RESEARCH CENTER
Paris - Rocquencourt

FIELD

Activity Report 2014

Section New Results

Edition: 2015-03-24

ANTIQUE Team

6. New Results

6.1. Highlights of the Year

Patrick and Radhia Cousot have received in 2014 the IEEE Computer Society IEEE Computer Society Harlan D. Mills award for the invention of abstract interpretation, developpment of tool support and practical application http://www.computer.org/portal/web/awards/cousots.

6.2. Memory Abstraction

6.2.1. Modular Construction of Shape-Numeric Analyzers

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

Abstract interpretation, Memory abstraction, Shape abstract domains. In [24], we discuss the modular construction of memory abstract domains.

The aim of static analysis is to infer invariants about programs that are tight enough to establish semantic properties, like the absence of run-time errors. In the last decades, several branches of the static analysis of imperative programs have made significant progress, such as in the inference of numeric invariants or the computation of data structures properties (using pointer abstractions or shape analyzers). Although simultaneous inference of shape-numeric invariants is often needed, this case is especially challenging and less well explored. Notably, simultaneous shape-numeric inference raises complex issues in the design of the static analyzer itself. We studied the modular construction of static analyzers, based on combinations of atomic abstract domains to describe several kinds of memory properties and value properties.

6.2.2. An abstract domain combinator for separately conjoining memory abstractions

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

Abstract interpretation, Memory abstraction, Shape abstract domains. In [25], we studied the separating combination of heap abstract domains.

The breadth and depth of heap properties that can be inferred by the union of today's shape analyses is quite astounding. Yet, achieving scalability while supporting a wide range of complex data structures in a generic way remains a long-standing challenge. We proposed a way to side-step this issue by defining a generic abstract domain combinator for combining memory abstractions on disjoint regions. In essence, our abstract domain construction is to the separating conjunction in separation logic as the reduced product construction is to classical, non-separating conjunction. This approach eases the design of the analysis as memory abstract domains can be re-used by applying our separating conjunction domain combinator. And more importantly, this combinator enables an analysis designer to easily create a combined domain that applies computationally-expensive abstract domains only where it is required.

6.2.3. Abstraction of Arrays Based on Non Contiguous Partitions

Participants: Xavier Rival [correspondant], Jiangchao Liu.

Abstract interpretation, Memory abstraction, Array abstract domains. In [20], we studied array abstractions.

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

6.3. Static analysis of JavaScript applications

6.3.1. Automatic Analysis of Open Objects in Dynamic Language Programs

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

Abstract interpretation, Dynamically typed languages, Verification In [14], we have studied the abstraction of open objects in dynamic language programs (like JavaScript).

In dynamic languages, objects are open: they support iteration over and dynamic addition/deletion of their attributes. Open objects, because they have an unbounded number of attributes, are difficult to abstract without a priori knowledge of all or nearly all of the attributes and thus pose a significant challenge for precise static analysis. To address this challenge, we presented the HOO (Heap with Open Objects) abstraction that can precisely represent and infer properties about open-object-manipulating programs without any knowledge of specific attributes. It achieves this by building upon a relational abstract domain for sets that is used to reason about partitions of object attributes. An implementation of the resulting static analysis is used to verify specifications for dynamic language framework code that makes extensive use of open objects, thus demonstrating the effectiveness of this approach.

6.3.2. Desynchronized Multi-State Abstractions for Open Programs in Dynamic Languages

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

Abstract interpretation, Dynamically typed languages, Verification In [15], we have studied desynchronized multi-state abstractions for open programs in dynamic languages (libraries).

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

6.4. Static analysis of Spreadsheet applications

Participants: Tie Cheng [correspondant], Xavier Rival.

Abstract interpretation, Spreadsheet applications, Verification In [13], we have proposed a static analysis to detect type unsafe operations in spreadsheet applications including formulas and macros.

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

6.5. Mechanically Verifying a Shape Analysis

Participant: Arnaud Spiwack.

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.

6.6. Static Analysis of Embedded Critical Concurrent Software

6.6.1. AstréeA: A Static Analyzer for Large Embedded Multi-Task Software

Participant: Antoine Miné.

In [11], we present the design, implementation and experimentation of the ASTRÉEA static analyzer, an extension of the ASTRÉE static analyzer dedicated to analyzing the run-time errors in embedded critical concurrent software. Such software are already present in critical systems and will likely become the norm with the generalization of multi-core processors in embedded systems, leading to new challenging demands in verification. One major challenge is that a concurrent program execution does not follow a fixed sequential order, but one of many interleavings of executions from different tasks chosen by the scheduler. As it is impractical to build a fully flow-sensitive analysis by enumerating explicitly all interleavings, we took inspiration from thread-modular methods: we analyze each thread individually, in an environment consisting of (an abstraction of) the effect of the other threads. This is a form of rely-guarantee reasoning, but in a fully automatic static analysis settings formalized as abstract interpretation: a thread-modular static analysis is viewed as a computable abstraction of a complete concrete, fixpoint-based thread-modular semantics. This permits a fine control between precision and efficiency, and opens the way to analysis specialization: any given safety property of a given program can be theoretically inferred given the right abstract domain. The presentation describes our subsequent work in improving the precision of ASTRÉEA by specialization on our target applications, and the interesting abstractions we developed along the way. For instance, we developed new interference abstractions enabling a limited but controllable (for efficiency) degree of relationality and flow-sensitivity. We also designed abstractions able to exploit our knowledge of the real-time scheduler used in the analysis target: i.e., it schedules tasks on a single core and obeys a strict priority scheme. The end-result is a more precise analyzer on our target applications, with currently around a thousand alarms.

6.6.2. Static Analysis by Abstract Interpretation of Concurrent Programs under the TSO Weak Memory Model

Participants: Thibault Suzanne, Antoine Miné.

In [33], we present an abstract semantics for the Total Store Ordering (TSO) memory model, a weakly consistent memory model used in major multi-core processors. This abstraction forgets some information about the order in which variables are written into by each thread. This results in a much simplified concrete semantics, but which is still not computable. We then express the semantics based on partitioned sets of points in a vector space, which allows applying classic methods from abstract interpretation (such as numeric abstract domains) to achieve a fully computable abstract semantics and automatically infer an over-approximation of the set of reachable states of a program running under the TSO memory model. The method is proved correct and, in certain cases, optimal, using the standard tools of abstraction interpretation (Galois connections).

Moreover, we have written a prototype static analyzer for simple program fragments written in an assembly-like language, and experimented our abstraction on a few small examples.

6.7. Inference of Termination and Liveness properties

6.7.1. A Decision Tree Abstract Domain for Proving Conditional Termination

Participants: Caterina Urban, Antoine Miné.

In [26], we present a new parameterized abstract domain able to refine existing numerical abstract domains with finite disjunctions. The elements of the abstract domain are decision trees where the decision nodes are labeled with linear constraints, and the leaf nodes belong to a numerical abstract domain. The abstract domain is parametric in the choice between the expressivity and the cost of the linear constraints for the decision nodes (e.g., polyhedral or octagonal constraints), and the choice of the abstract domain for the leaf nodes. We describe an instance of this domain based on piecewise-defined ranking functions for the automatic inference of sufficient preconditions for program termination. We have implemented a static analyzer for proving conditional termination of programs written in (a subset of) C and, using experimental evidence, we show that it performs well on a wide variety of benchmarks, it is competitive with the state of the art and is able to analyze programs that are out of the reach of existing methods.

6.7.2. An Abstract Domain to Infer Ordinal-Valued Ranking Functions

Participants: Caterina Urban, Antoine Miné.

The traditional method for proving program termination consists in inferring a ranking function. In many cases (i.e. programs with unbounded non-determinism), a single ranking function over natural numbers is not sufficient. In [30], we propose a new abstract domain to automatically infer ranking functions over ordinals. We extend an existing domain for piecewise-defined natural-valued ranking functions 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 state partitioning inducing the piecewise-definition and the type of functions used as polynomial coefficients. 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 allows us to take a first step in the direction of proving termination under fairness constraints and proving liveness properties of (sequential and) concurrent programs.

6.7.3. Proving Guarantee and Recurrence Temporal Properties by Abstract Interpretation

Participants: Caterina Urban, Antoine Miné.

We present in [28] a new static analysis methods for proving liveness properties of programs. In particular, with reference to the hierarchy of temporal properties proposed by Manna and Pnueli, we focus on guarantee (i.e., “something good occurs at least once”) and recurrence (i.e., “something good occurs infinitely often”) temporal properties. We generalize the abstract interpretation framework for termination presented by Cousot and Cousot. Specifically, static analyses of guarantee and recurrence temporal properties are systematically derived by abstraction of the program operational trace semantics. These methods automatically infer sufficient preconditions for the temporal properties by reusing existing numerical abstract domains based on piecewise-defined ranking functions. We augment these abstract domains with new abstract operators, including a dual widening. To illustrate the potential of the proposed methods, we have implemented a research prototype static analyzer, for programs written in a C-like syntax, that yielded interesting preliminary results.

6.8. Numeric Invariant Inference

6.8.1. A Numeric Abstract Domain to Infer Octagonal Constraints with Absolute Value

Participants: Liqian Chen [National Laboratory for Parallel and Distributed Processing, National University of Defense Technology, Changsha, P.R.China], Jiangchao Liu, Antoine Miné, Deepak Kapur [University of New Mexico, USA], Ji Wang [National Laboratory for Parallel and Distributed Processing, National University of Defense Technology, Changsha, P.R.China].

The octagon abstract domain, devoted to discovering octagonal constraints (also called Unit Two Variable Per Inequality or UTVPI constraints) of a program, is one of the most commonly used numerical abstractions in practice, due to its quadratic memory complexity and cubic time complexity. However, the octagon domain itself is restricted to express convex sets and has limitations in handling non-convex properties which are sometimes required for proving some numerical properties in a program. In [12], we intend to extend the octagon abstract domain with absolute value, to infer certain non-convex properties by exploiting the absolute value function. More precisely, the new domain can infer relations of the form $\{ \pm X \pm Y \leq c,\ \pm X \pm |Y| \leq d,\ \pm|X| \pm |Y| \leq e \}$. We provide algorithms for domain operations such that the new domain still enjoys the same asymptotic complexity as the octagon domain. Moreover, we present an approach to support strict inequalities over rational or real-valued variables in this domain, which also fits for the octagon domain. Experimental results of our prototype are encouraging; The new domain is scalable and able to find non-convex invariants of interest in practice but without too much overhead (compared with that using octagons).

6.8.2. *A Method to Infer Inductive Numeric Invariants Inspired from Constraint Programming.*

Participant: Antoine Miné.

In [29], we suggest the idea of using algorithms inspired by Constraint Programming in order to infer inductive invariants on numeric programs. Similarly to Constraint Programming solvers on continuous domains, our algorithm approximates the problem from above, using decreasing iterations that may split, discard, and tighten axis-aligned boxes. If successful, the algorithm outputs a set of boxes that includes the initial states and is a post-fixpoint of the abstract semantic function of interest. Our work is very preliminary; many improvements still need to be performed to determine if the method is usable in practice, and in which contexts. Nevertheless, we show that a naive proof-of-concept implementation of our algorithm is already capable of inferring non-trivial inductive invariants that would otherwise require the use of relational or even non-linear abstract domains when using more traditional abstract interpretation iteration methods.

6.9. Bisimulation Metrics

6.9.1. *Bisimulation for Markov Decision Processes through Families of Functional Expressions*

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

Markov decision processes, Bisimulation, Metrics.

In [17], we have transfered a notion of quantitative bisimilarity for labelled Markov processes [51] 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 [56], [55] 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 [52]; what is novel is that we recast this proof in terms of integral probability metrics [54] 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.

6.9.2. *Bisimulation Metrics are Optimal Value Functions*

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

Markov decision processes, Bisimulation, Metrics.

Bisimulation is a notion of behavioural equivalence on the states of a transition system. Its definition has been extended to Markov decision processes, where it can be used to aggregate states. A bisimulation metric is a quantitative analog of bisimulation that measures how similar states are from a the perspective of long-term behavior. Bisimulation metrics have been used to establish approximation bounds for state aggregation and other forms of value function approximation. In [18], we prove that a bisimulation metric defined on the state space of a Markov decision process is the optimal value function of an optimal coupling of two copies of the original model. We prove the result in the general case of continuous state spaces. This result has important implications in understanding the complexity of computing such metrics, and opens up the possibility of more efficient computational methods.

6.10. Abstraction of Rule-Based Biological Models

6.10.1. Stochastic fragments: A framework for the exact reduction of the stochastic semantics of rule-based models

Participants: Jérôme Feret, Heinz Koeppl [École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland], Tatjana Petrov [École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland].

Protein-protein interaction networks, Stochastic systems, Backward bisimulations, Model reduction. In [9], we propose an abstract interpretation-based framework for reducing the state space of stochastic semantics for protein-protein interaction networks. Our approach consists in quotienting the state space of networks. Yet interestingly, we do not apply the widely-used strong lumpability criterion which imposes that two equivalent states behave similarly with respect to the quotient, but a weak version of it. More precisely, our framework detects and proves some invariants about the dynamics of the system: indeed the quotient of the state space is such that the probability of being in a given state knowing that this state is in a given equivalence class, is an invariant of the semantics. Then we introduce an individual-based stochastic semantics (where each agent is identified by a unique identifier) for the programs of a rule-based language (namely Kappa) and we use our abstraction framework for deriving a sound population-based semantics and a sound fragments-based semantics, which give the distribution of the traces respectively for the number of instances of molecular species and for the number of instances of partially defined molecular species. These partially defined species are chosen automatically thanks to a dependency analysis which is also described in [9].

6.10.2. An algebraic approach for inferring and using symmetries in rule-based models

Participant: Jérôme Feret.

Graph rewriting, Single-pushout semantics, Symmetries, Bisimulations, Model reduction. Symmetries arise naturally in rule-based models, and under various forms. Besides automorphisms between site graphs, which are usually built within the semantics, symmetries can take the form of pairs of sites having the same capabilities of interactions, of some protein variants behaving exactly the same way, or of some linear, planar, or 3D molecular complexes which could be seen modulo permutations of their axis and/or mirror-image symmetries. In [16], we propose a unifying handling of symmetries in Kappa. We follow an algebraic approach, that is based on the single pushout semantics of Kappa. We model classes of symmetries as finite groups of transformations between site graphs, which are compatible with the notion of embedding (that is to say that it is always possible to restrict a symmetry that is applied with the image of an embedding to the domain of this embedding) and we provide some assumptions that ensure that symmetries are compatible with pushouts. Then, we characterise when a set of rules is symmetric with respect to a group of symmetries and, in such a case, we give sufficient conditions so that this group of symmetries induces a forward bisimulation and/or a backward bisimulation over the population semantics.

6.11. Model checking of Logical Biological Models

6.11.1. Model checking logical regulatory networks

Participants: Pedro T. Monteiro [INESC-ID, Lisboa, Portugal], Wassim Abou-Jaoudé, Denis Thieffry [IBENS, France], Claudine Chaouiya [IGC, Oeiras, Portugal].

Model checking, Regulatory networks. Regulatory and signalling networks control cell behaviours in response to environmental cues. The logical formalism has been widely employed to study these interaction networks, which are modelled as discrete dynamical systems. While biologists identify networks encompassing more and more components, properties of biological relevance become hard to verify.

In [22], we report on the use of model-checking techniques to address this challenge. This approach is illustrated by an application dealing with the modelling of T-helper lymphocyte differentiation.

6.11.2. Model checking to assess T-helper cell plasticity

Participants: Wassim Abou-Jaoudé, Pedro T. Monteiro [INESC-ID, Lisboa, Portugal], Aurélien Naldi [Centre Intégratif de Lausanne, Lausanne, Switzerland], Maximilien Grandclaudon [Institut Curie, Paris, France], Vassili Sommeils [Institut Curie, Paris, France], Claudine Chaouiya [IGC, Oeiras, Portugal], Denis Thieffry [IBENS, France].

Model checking, Logical modeling. Computational modeling constitutes a crucial step toward the functional understanding of complex cellular networks. In particular, logical modeling has proven suitable for the dynamical analysis of large signaling and transcriptional regulatory networks. In this context, signaling input components are generally meant to convey external stimuli, or environmental cues. In response to such external signals, cells acquire specific gene expression patterns modeled in terms of attractors (e.g., stable states).The capacity for cells to alter or reprogram their differentiated states upon changes in environmental conditions is referred to as cell plasticity. In this article, we present a multivalued logical framework along with computational methods recently developed to efficiently analyze large models.We mainly focus on a symbolic model checking approach to investigate switches between attractors subsequent to changes of input conditions. As a case study, we consider the cellular network regulating the differentiation of T-helper (Th) cells, which orchestrate many physiological and pathological immune responses. To account for novel cellular subtypes, we present, in [8], an extended version of a published model of Th cell differentiation.We then use symbolic model checking to analyze reachability properties between Th subtypes upon changes of environmental cues. This allows for the construction of a synthetic view of Th cell plasticity in terms of a graph connecting subtypes with arcs labeled by input conditions. Finally, we explore novel strategies enabling specific Th cell polarizing or reprograming events.

AOSTE Project-Team

6. New Results

6.1. Languages, Models of Computation and Metamodeling using logical clock constraints

Participants: Julien Deantoni, Robert de Simone, Frédéric Mallet, Marie Agnès Peraldi Frati.

A revised and updated version of our previous work on UML MARTE Time Model was written in survey textbook form for a larger audience, and published in [38]. Same was done for the more applied specific findings of the ARTEMIS PRESTO European project [39]. Also, a research report finalizing the denotational semantics of the logical clock constraint languages was issued for reference [44].

6.2. Experiments with Architecture and Application modeling

Participants: Robert de Simone, Émilien Kofman, Jean-Vivien Millo, Amine Oueslati, Mohamed Bergach.

We submitted for publication our theoretical results on formal mapping of an application written as a process network dataflow graph onto an abstract architecture model involving a network-on-chip and manycore processor arrays [24].

In the context of the *FUI Clistine* collaborative project (which aims at building a cheap supercomputer by assembling low-cost, general-purpose and network processors interconnected by a time-predictable, on-board network), we considered the issue of classifying general application types, in the fashion inherited from UC. Berkeley's 13 "dwarfs" [46]. Meanwhile,the modeling of desired architecture was slightly postponed due to hesitations from the main industrial partner (that will build the prototype itself). This work was the topic of Amine Oueslati's first year PhD. The classification, and the use of distinct type properties for efficient and natural encoding, was applied on typical application programs provided by partners (Galerkin methods for electromagnetic simulation by the Nachos Inria team, ray-tracing algorithms by the Optis/Simplysim SME design company).

In the context of Mohammed Bergach's CIFRE PhD contract with Kontron Toulon, we conducted an advanced modeling exercice on how to best fit large DFT (Discrete Fourier Transform) modules onto a specific processor architecture (first Intel Sandybridge, then Haswell) that offers computing compromise costs (in performance vs power) between regular CPUs and GPU hardware accelerators. There were two issues: first, how to best dimension the size of the largest FFT block that may be performed locally on a corresponding GPU compute block; second, how to distribute the many such optimal size FFT block needed in a typical radar application, using the GPU and CPU features at the best of their capacity, with account of the slow data transfer latencies across memory banks (to and from the GPU registers).
As a side-effect, people from Kontron are now using and distributing to their customers the FFT GPU libraries with ad-hoc FFT variants matching the GPU block memory sizes. The development, rather lenghty in the case of Sandybridge, was quickly adjusted and ported for Haswell. A new workshop paper is under submission.

6.3. Multiview modeling with performance and power aspects

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

In the context of the ANR HOPE project and The CIM PACA Design Platform, we continued our work on joint modeling and co-simulation of abstract architecture and application (use case scenario) models, together with non-functional aspect views such as performance, power and temperature. The goal of the HOPE project is to consider *hierarchical* {Power/Performance/Temperature} Management Units (MU), and our target is to connect our IDM modeling with dedicated tools such as Synopsys Platform Architect or Docea Power AcePlorer. The IP_Xact interface format for IP blocks is also aimed for compositional assembly representation, including non-functional properties and timing semantics constraints for co-simulation. This work is mostly continued from the former PhD thesis of Carlos Gomez to a new framework by Ameni Khecharem, as part of her PhD. Practical co-simulation trends are also investigated. Currents results were reported in [29]

6.4. Heterogeneous Languages Coordination with Concurrency and Time

Participants: Julien Deantoni, Matias Vara Larsen, Robert de Simone, Frédéric Mallet.

In the context of the ANR GEMOC project and in closely related to the mutiview approach of the team, we focused on how to deal with analysis and simulation of heterogeneous languages. Supporting coordinated use of heterogeneous domain specific languages leads to what we called the globalization of modeling language [22]. Concretely, we proposed to define a language behavioral interface to exhibit the concurrency and time aspects of the semantics of a language. The concurrency and time aspects are described by a formal extension of CCSL, named MoCCML (Model of Concurrency and Communication Modeling Language [45], [28]). Any models that conform such language exhibit a symbolic representation of all its acceptable schedules. Based on this, we shown that it is possible to coordinate heterogeneous models .To avoid redundant model coordination activities, we reified the know-how about model coordination in BCOoL (Behavioral Coordination Operator Language[34]), a language dedicated to language coordination. This work is mainly realized by Matias Vara Larsen, as part of his PhD. In this context, we organize the community around such subject for the second year in an international workshop [43] with an increasing number of participants.

6.5. Performance study of Massively Parallel Processor Array (MPPA) SoC architecture

Participants: Sid Touati, Franco Pestarini.

From a previous collaboration programme, we (Aoste Sophia) possess a MMPA manycore chip, designed and produced by the company Kalray, in Grenoble. The chip integrates 256 cores, composed of 16 clusters (themselves each with 16 cores), and a powerfull network-on-chip interconnect mesh structure. This architecture is oriented towards high performance embedded application, with real time constraints. The cores and NoC were designed to deliver predictable performance.

Our current project, during Franco Pestarini Inria International Intern period, was to test the performance of the NoC, trying to obtain better knowledge of its behavior. We put up a set of microbenchmarks to exercice the network under different specific scenarios (low overhead network traffic, high traffic), and analyzed the experimental results.

We produced a detailed deliverable report explaining under which conditions the NoC could deliver stable and predictable performances. We identified potential configurations where the network becomes unstable (leading to variable and impredictable performances and bandwith).

Meanwhile, the textbook on low-level code optimization, written between Sid Touati and Kalray CTO, appeared in published form [42]. Its content reports on some of the techniques used inside the MPPA compilation environment, and beyond.

6.6. Parametric and Non Parametric Statistics for Code Performance Analysis

Participant: Sid Touati.

This activity is conducted by Sid Touati in collaboration with Julien Worms, an associate professor in Mathematics at the university of Versailles Saint Quentin. It was started under the consideration that the performances of programs are hardly ever represented by a gaussian distribution. So, our purpose here is to study parametric statistics for analyzing the performances of programs. We are interested in modelling program performances by gaussian mixtures (using mixmod method). After a statistical modeling, we deduce multiple performance tests and performance criteria to decide with a high degree of confidence about the "best" program run version. This is still work-in-progress: we are implementing a free software for analysis based on our approach, and we are writing a rather complete research report prior to further publications in conferences and journals.

6.7. Uniprocessor Real-Time Scheduling

Participants: Falou Ndoye, Yves Sorel, Walid Talaboulma.

6.7.1. Real-Time Scheduling with Exact Preemption Cost

Previous years, we worked on schedulability analyses of dependent tasks, executed on a uniprocessor, which take into account the exact preemption cost and more generally the cost of the real-time operating system. Indeed, this cost is composed of the cost of the preemptions and the cost of the scheduler. Our approach is based on an offline real-time schedulability analysis, proved sustainable, that produces a scheduling table. This latter contains the next instants (activation and completion of tasks) when the scheduler will be called, beeing aware of the instants where tasks are preempted and then resumed. This approach allows the schedulability analysis to account preemption costs involved by other preemptions. The scheduling table contains also the address of the next task to execute preventing the scheduler to choose it in the ready tasks list, unlike with classical on-line scheduler. The theoretical results in the uniprocessor case, are given in the Falou Ndoye's PhD thesis [19] defended in April this year. This approach has been implemented through an offline scheduler that is triggered by a timer when this latter is equal to zero and loaded with the next instant contained in the scheduling table, according to a time trigger approach.

We carried out two kinds of implementations. Actually, the first one is a simulation since our time trigger offline scheduler is modelled as a high priority task running uppon an existing operating system. We experimented this approach with Vanilla Linux (not modified) and Linux/Xenomai, real-time versions of the Linux operating system, with low latency caracteristics. Of course, these implementations were only able to show that the theoterical results were correct, but did not provide good real-time performances nor a robust way to measure time without influencing the usefull code. Therefore, we implemented our scheduler on a bare metal ARM968E-S processor based on an ARM9 architecture since it is widespread in the industry world, and we experimented this processor few years ago to determine the cost of classical online schedulers.

For this purpose we used a MCB2929 developpement board, from Keil, containing the LPC2929 SoC including the ARM968E-S, an accurate timer, and various peripherals. The scheduling table is generated offline for a set of tasks, and stored in the memory as an array of couples, each composed of the task to execute and the duration elapsing until the next scheduler call. This duration is used to set the timer counter. When it hits zero it triggers a high priority interruption that is serviced by calling again the scheduler to choose the next couple (task, duration), and so on up to the end of the scheduling table. This will repeat infinitely from the beginning of the scheduling table.

We tested different set of tasks with multiple preemption scenarios, that can yield to deadlines misses. We measured for a **12Mhz** CPU and timer clock frequencies a value of **28** μ**s** for the scheduler cost, and of **1** μ**s** for the preemption cost.

6.8. Multiprocessor Real-Time Scheduling

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

6.8.1. Multiprocessor Partitioned Scheduling with Exact Preemption Cost

Since we chose a multiprocessor partitioned scheduling approach, we can take advantage of the results we obtained in the case of uniprocessor real-time Scheduling accounting for the cost of the real-time operating system, i.e. the cost of preemptions and of the scheduler. From the point of view of the off-line real-time schedulability analysis we only have to consider in addition to activation and completion instants, reception of data instants. This latter instant is determinated by supposing that the cost of every data transfer is known for every possible communication medium. Indeed, when two dependent tasks are allocated to two different processors, the consuming task will have to wait for the data sent by the producing task. The theoretical results in the multiprocessor case, are given in the Falou Ndoye's PhD thesis [19] defended in April this year. We chose the message passing protocol for interprocessor communications achieved through a switched ethernet network. In order to determine precisely the cost of data transfers, we started to investigate the possible approaches to synchronize the send and receive tasks located in two different processors and to schedule them with the other tasks allocated to the same processor. This synchronization protocol will be taken into account to determine the interprocessor communication costs. Concerning these communication costs, we consider FIFO and FIFO* schedulings in the switches, the later is a FIFO scheduling based on the release time of frames at their source node. We have first corrected the trajectory approach (recently shown to be optimistic for corner cases) with FIFO scheduling to compute worst case end-to-end communication costs. Then, we have extended the trajectory approach to FIFO* scheduling. We want to implement our off-line scheduler on every processor of a multiprocessor architecture composed of at least three processors, communicating through an ethernet switch.

Concerning the delay of communications, we consider FIFO and FIFO* schedulings in the switches, the later is a FIFO scheduling based on the release time of frames at their source node. We have first corrected the trajectory approach (recently shown to be optimistic for corner cases) with FIFO scheduling to computed worst case end-to-end communication delays. Then, we have extended the trajectory approach to FIFO* scheduling.

6.8.2. Mutiprocessor Parallel Directed Acyclic Graph (DAG) scheduling

We are interested in studying the hard real-time scheduling problem of parallel Directed Acyclic Graph (DAG) tasks on multiprocessor systems. In this model, a task is defined as a set of dependent subtasks that execute under precedence constraints. The execution order of these subtasks is dynamic, i.e., a subtask can execute either sequentially or in parallel with its siblings based on the decisions of the real-time scheduler. To this end, we analyze two DAG scheduling approaches to determine the execution order of subtasks: the Model Transformation and the Direct Scheduling approaches. We consider global preemptive multiprocessor scheduling algorithms to be used with the scheduling approaches, such as Earliest Deadline First (EDF) and Deadline Monotonic (DM).

6.8.3. Gateway with Modeling Languages for Certified Code Generation

This work was carried out in the P FUI project 8.2.2 We continued the work on the gateway between the P formalism and SynDEx, started the last two years. We have integrated in the gateway the IF and FOR blocks of Simulink that were missing in the functional specification, except for particular cases where the IF block is nested in the FOR block, or the opposite. The integration of the MERGE and MUX blocks are still to be done. We extended the P formalism with architectural elements that SynDEx needs to perform schedulability analyses on functional specifications. These architectural elements are hardware resources (processor, bus, shared memory, router) and timing characteristics (deadline, period, WCET, WCTT). We developed a new part in the gateway which transforms an architectural model described with the P formalism in the input format of SynDEx. We developed also a third part in the gateway which feedbacks the schedulability analysis results obtained with SynDEx (the scheduling table) and stores them into models described with the P formalism. Finally, we have collaborated with the industrial partners to test our gateway on their use cases.

6.8.4. SynDEx updates

The first tests on the alpha version of SynDEx V8, released last year, shown some bugs that we fixed. This first release did not include a code generator. Thus, we worked to interface the distributed real-time embedded code generator of SynDEx V7 with SynDEx V8.

6.9. Probabilistic Real-Time Systems

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

The advent of 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 [23] we have described the vision of FP7 IP PROXIMA, project that is interested in the introduction of randomization of the architectures at cache level. For this type of architecture static probabilistic timing analysis is possible [20] by providing bounds on the probabilistic execution time of a task. An industrial case study from avionics is detailed in [32]. Such distribution is then used as input for probabilistic scheduling as described in [37], [36].

This year we have also provided a complete state of the art of the probabilistic real-time systems in [17].

6.10. Off-line (static) mapping and WCET analysis of real-time applications onto NoC-based many-cores

Participants: Dumitru Potop Butucaru, Thomas Carle, Manel Djemal, 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 have 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. Furthermore, if in addition the computation and memory resources of the MPSoC have support for real-time predictability, then parallel applications can be developed that allow very precise WCET analysis of parallel code. WCET analysis of parallel code is joint work with Isabelle Puaut of Inria, EPI ALF.

This year we have completed our mapping (allocation and scheduling) and code generation technique and tool for NoC-based MPSoCs. NoCs pose significant challenges to both on-line (dynamic) and off-line (static) real-time scheduling approaches [25]. 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, efficient resource allocation requires scalable algorithms working on hardware models with a level of detail that is unprecedented in real-time scheduling.

We considered a static (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[26] and pipelined scheduling [21].

Our method has been implemented within the Lopht tool presented in section 5.4 , and first results are presented in [26], [25], and in extenso in the PhD thesis of manel Djemal [18]. One of the objectives of the starting CAPACITES project is the evaluation of the possibility of porting Lopht and the WCET analysis technique for parallel code onto the Kalray MPPA platform.

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

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

We have continued this year the work on real-time scheduling and code generation for time-triggered platforms. Much of this work was carried out as part of a bilateral collaboration with Airbus DS and the CNES, which fund the post-doctorate of Raul Gorcitz, and in our collaboration with the IRT SystemX, project FSF.

The objective is to facilitate the development of complex time-triggered systems by automating the allocation, scheduling, and code generation steps. We show that full automation is possible while taking into account all the specification elements required by a complex, real-life embedded control system. The main originality of our work is that it takes into account at the same time multiple complexity elements: functional specifications with conditional execution and multiple modes and various types of non-functional properties: real-time (release dates, deadlines, major time frame, end-to-end flows), ARINC 653 partitioning (which we can fully or partially synthesize), task preemptability, allocation. Our algorithms allow the automatic allocation and scheduling onto multi-processor (distributed) systems with a global time base, taking into account communication costs.

While the past years were mainly dedicated to the development of this scheduling and code generation technique, this year the technique and the associated tool have matured enough to allow the publication of the first results concerning the optimized scheduling algorithms [21] and its application on large case studies. Ongoing work by the post-doc Raul Gorcitz, funded by Airbus DS and the CNES aims at evaluating the applicability of our methods on embedded platforms that are being considered for the future european space launchers. The Lopht tool is also used in the IRT SystemX, project FSF as part of the proposed design flow. All extensions have been implemented in the Lopht tool. All this work has been presented *in extenso* in the PhD thesis of Thomas Carle [16].

CASCADE Project-Team

5. New Results

5.1. Results

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

- More efficient constructions with lattices
- New constructions from pairings
- Delegation of computations
- Analysis of pseudo-random generators
- Advanced primitives for the privacy in the cloud
- Cryptanalysis of symmetric primitives
- New leakage-resilient primitives
- Stronger security with related-key security

CRYPT Team

4. New Results

4.1. Highlights of the Year

The team published [20] improved single-key attacks on reduced-round AES: AES is currently the most widespread block cipher standard, it is implemented in Intel processors.

The team also showed [18] how to speed-up a well-known public-key cryptanalysis technique: finding small roots of univariate polynomial congruences. This technique is used to break special cases of the RSA cryptosystem.

Phong Nguyen was Program co-Chair of the 33rd IACR Eurocrypt Conference (EUROCRYPT 2014) [22].

DEDUCTEAM Exploratory Action

6. New Results

6.1. Highlights of the Year

In the framework of the *BWare* project, Pierre Halmagrand, David Delahaye, Damien Doligez, and Olivier Hermant designed a new version of the *B* set theory using deduction modulo, in order to automatically verify a large part of the proof obligations of the benchmark of *BWare*, which consists of proof obligations coming from the modeling of industrial applications (about 13,000 proof obligations). Using this *B* set theory modulo with *Zenon Modulo*, as well as some other extensions of *Zenon*, such as typed proof search and arithmetic (implemented by Guillaume Bury), we are able to automatically verify more than 95% of the proof obligations of *BWare*, while the regular version of *Zenon* is only able to prove less than 1% of these proof obligations. This is a real breakthrough for the *BWare* project, but also for automated deduction in general, as it tends to show that deduction modulo is the way to go when reasoning modulo theories.

6.2. Termination

Frédéric Blanqui, together with Jean-Pierre Jouannaud (Univ. Paris 11) and Albert Rubio (Technical University of Catalonia), have finished their work on a new version of the higher-order recursive path ordering (HORPO) [44], [43], a decidable monotone well-founded relation that can be used for proving the termination of higher-order rewrite systems by checking that rules are included in it. This new version, called the computability path ordering (CPO), appears to be the ultimate improvement of HORPO in the sense that this definition captures the essence of computability arguments *à la* Tait and Girard [37], 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 new result is described in a 41-pages papers available on Frédéric Blanqui's web page which has been submitted to a journal for publication. A Prolog implementation of CPO is also available on Albert Rubio's web page.

Frédéric Blanqui revised his work on the compatibility of Tait and Girard's notion of computability for proving the termination of higher-order rewrite systems when matching is done modulo $\beta\eta$-equivalence. In particular, he showed that computability is preserved by leaf-β-expansion, a key property for dealing with higher-order pattern-matching. This work is described in a 46-pages paper available on his web page which has been submitted to a journal for publication.

Frédéric Blanqui did some historical investigations on fixpoint theorems in posets used for instance for defining the semantics of non-basic inductive types (i.e. types with constructors taking functions as arguments) and the termination of functions defined by induction on such non-basic inductive types. These theorems assume the function either extensive or monotone. However, as shown by Salinas in [48], these two conditions can be subsumed by a more general one. Frédéric Blanqui slightly improved this condition further by using results by Hartogs, Rubin and Rubin, and Abian and Brown. This work is described in a 10-pages note available on his web page [20].

Kim Quyen Ly finished the development of a new version faster, safer (proved correct in Coq) and standalone version of Rainbow, based on Coq extraction mechanism. She defended her PhD thesis [11] on the automated verification of termination certificates in October.

6.3. Proof and type theory modulo rewriting

Ali Assaf defined a sound and complete embedding of the cumulative universe hierarchy of the *calculus of inductive constructions* (CIC) in the $\lambda\Pi$-calculus modulo rewriting [18]. By reformulating universes in the Tarski style, he showed that we can make cumulativity explicit without losing any typing power. This result refines the translation used by Coqine, which was unsound because it collapsed the universe hierarchy to a single type universe. It also sheds some light on the metatheory of Coq and its connection to Martin-Löf's intuitionistic type theory. This work was presented at the TYPES meeting in Paris.

Frédéric Gilbert and Olivier Hermant defined new encodings from classical to intuitionistic first-order logic. These encodings, based on the introduction of double negations in formulas, are tuned to satisfy two purposes jointly: basing their specifications on the definition of *classical connectives* inside intuitionistic logic – which is the property of *morphisms*, and reducing their impact on the shape and size of formulas, by limiting as much as possible the number of negations introduced. This paper has been submitted.

Raphael Cauderlier and Catherine Dubois defined a shallow embedding of an object calculus (formalized by Abadi and Cardelli), in the $\lambda\Pi$-calculus modulo rewriting . The main result concerns the encoding of subtyping. This encoding shows that rewriting is an effective help for handling of subtyping proofs. The implementation in Dedukti, Sigmaid. This work has been presented at the TYPES 2014 meeting in Paris. A paper has been submitted.

Ali Assaf, Olivier Hermant and Ronan Saillard defined a rewrite system such that all strongly normalizable proof term can be typed in Natural Deduction modulo this rewrite system. This work is inspired by Statman's work [49], and can be understood as an encoding of intersection types.

Guillaume Burel showed how to get rewriting systems that admit cut by using standard saturation techniques from automated theorem proving, namely ordered resolution with selection, and superposition. This work relies on a view of proposition rewriting rules as oriented clauses, like term rewriting rules can be seen as oriented equations. This also lead to introduce an extension of deduction modulo with *conditional* term rewriting rules. This work was presented at the RTA-TLCA conference in Vienna [15].

Gilles Dowek, has generalized the notion of super-consistency to the lambda-Pi-calculus modulo theory and proved this way the termination of the embedding of various formulations of Simple Type Theory and of the Calculus of Constructions in the Lambda-Pi calculs modulo theory.

Gilles Dowek and Alejandro Díaz-Caro have finished their work on the extension of Simply Typed Lambda-Calculus with Type Isomorphisms. This work has been presented at the Types meeting and recently accepted for publication in the Theoretical Computer Science journal [26].

Gilles Dowek and Ying Jiang have given a new proof of the decidability of reachability in alternating pushdown systems, based on a cut-elimination theorem.

Vaston Costa presented to the group a new structure to represent proofs through references rather than copy. The structure, called Mimp-graph, was initially developed for minimal propositional logic but the results have been extended to first-order logic. Mimp-graph preserves the ability to represent any Natural Deduction proof and its minimal formula representation is a key feature of the mimp-graph structure, it is easy to distinguish maximal formulas and an upper bound in the length of the reduction sequence to obtain a normal proof. Thus a normalization theorem can be proved by counting the number of maximal formulas in the original derivation. The strong normalization follow as a direct consequence of such normalization, since that any reduction decreases the corresponding measures of derivation complexity. Sharing for inference rules is performed during the process of construction of the graph. This feature is very important, since we intend to use this graph in automatic theorem provers.

6.4. Automated theorem proving

Guillaume Bury defined a sound and complete extension of the tableaux method to handle linear arithmetic. The rules are based on a variant of the simplex algorithm for rational and real linear arithmetic, and a Branch&Bound algorithm for integer arithmetic.

Guillaume Bury defined an encoding of analytical tableaux rules as a theory for smt solvers. The theory acts like a lazy cnf conversion during the proof search and allows to integrate the cnf conversion into the resolution proof for unsatisfiable formulas. This work was implemented in mSAT.

Simon Cruanes added many improvements to Logtk, in particular a better algorithm to reduce formulas to Clausal Normal Form. A presentation of its design and implementation has been made at PAAR 2014[16]. He also used Zipperposition as a testbed for integer linear arithmetic; a sophisticated inference system for this fragment of arithmetic was designed and implemented in Zipperposition, including many redundancy criteria and simplification rules that make it efficient in practice. The arithmetic-enabled Zipperposition version entered CASC-J7, the annual competition of Automated Theorem Provers, in the first-order theorems with linear arithmetic division where it had very promising results (on integer problems only, since Zipperposition doesn't handle rationals).

Another extension of Zipperposition has been performed by Julien Rateau, Simon Cruanes, and David Delahaye, in order to deal with a fragment of set theory in the same vein as the *STR*$\dot{+}$*VE*$\subseteq$ prover [40]. This extension relies on a specific normal form of literal, which only involves the $\subseteq$, $\cap$, $\cup$, and complement set operators. In the future, the idea is to use this extension in the framework of the *BWare* project to verify *B* proof obligations coming from industrial benchmarks.

The current effort of research on Zipperposition focuses on extending superposition to handle structural induction, following the work from [45]. The current prototype is able to prove simple properties on natural numbers, binary trees and lists.

Kailiang Ji defined a set of rewrite rules for the equivalence between CTL formulas (denote them as R_{CTL}), by taking them as terms of designed predicates. For a given transition system model, we transform it into a set of rewrite rules (denote them as R_m). Then any CTL property of the transition system can be proved in deduction modulo $R_{CTL} \cup R_m$, by specifying the model checking problems into designed first-order formulas. This method was implemented in iProver Modulo, and the experimental evaluation was reported in workshop of Locali 2014.

6.5. Algebraic λ-calculus

Ali Assaf, Alejandro Díaz-Caro, Simon Perdrix, Christine Tasson, and Benoit Valiron completed a journal paper covering results on different algebraic extensions of the λ-calculus [12]. These extensions equip the calculus with an additive and a scalar-multiplicative structure, and their set of terms is closed under linear combinations. Two such extensions, the *algebraic λ-calculus* and the *linear-algebraic λ-calculus* arise independently in different contexts – the former is a fragment of the differential λ-calculus, the latter is a candidate λ-calculus for quantum computation – and have different operational semantics. In this paper, the authors showed how the two approaches relate to each other. They showed that the the first calculus follows a call-by-name strategy while the second follows a call-by-value strategy. They proved that the two can simulate each other using algebraic extensions of *continuation passing style* (CPS) translations that are sound and complete.

GALLIUM Project-Team

6. New Results

6.1. Formal verification of compilers and static analyzers

6.1.1. Formal verification of static analyzers based on abstract interpretation

Participants: Jacques-Henri Jourdan, Xavier Leroy, Sandrine Blazy [EPI Celtique], Vincent Laporte [EPI Celtique], David Pichardie [EPI Celtique], Sylvain Boulmé [Grenoble INP, VERIMAG], Alexis Fouilhe [Université Joseph Fourier de Grenoble, VERIMAG], Michaël Périn [Université Joseph Fourier de Grenoble, VERIMAG].

In the context of the ANR Verasco project, we are investigating the formal specification and verification in Coq of a realistic static analyzer based on abstract interpretation. This static analyzer handles a large subset of the C language (the same subset as the CompCert compiler, minus recursion and dynamic allocation); supports a combination of abstract domains, including relational domains; and should produce usable alarms. The long-term goal is to obtain a static analyzer that can be used to prove safety properties of real-world embedded C codes.

This year, Jacques-Henri Jourdan continued the developpment of this static analyzer. He finished the proof of correctness of the abstract interpreter, using an axiomatic semantics for the C#minor intermediate language to decompose this proof in two manageable halves. He improved the precision and performance of the abstract iterator and of numerical abstract domains. He designed and verified a symbolic domain that helps analyzing sequential Boolean operators such as `&&` and `||` that are encoded as Boolean variables and conditional constructs in the C#minor intermediate language. As a more flexible alternative to reduced products of domains, Jacques-Henri Jourdan designed, implemented and proved correct a communication system between numerical abstract domains, based on communication channels and inspired by Astrée [56].

In parallel, IRISA and VERIMAG, our academic partners on the Verasco project, contributed a verified abstract domain for memory states and pointer values (Vincent Laporte, Sandrine Blazy, and David Pichardie) and a polyhedric abstract domain for linear numerical inequalities (Alexis Fouilhe, Sylvain Boulmé, Michaël Périn) that uses validation a posteriori. Those various components were brought together by Jacques-Henri Jourdan and Vincent Laporte, resulting in an executable static analyzer.

The overall architecture and specification of Verasco is described in a paper [29] accepted for presentation at the forthcoming POPL 2015 conference.

6.1.2. The CompCert formally-verified compiler

Participants: Xavier Leroy, Jacques-Henri Jourdan.

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 [5]. This compiler comprises a back-end part, translating the Cminor intermediate language to PowerPC assembly and reusable for source languages other than C [4], 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 improved the CompCert C compiler in several directions:

- The parser, previously compiled to unverified OCaml code, was replaced by a parser compiled to Coq code and validated *a posteriori* by a validator written and proved sound in Coq. This validation step, performed when the CompCert compiler is compiled, provides a formal proof that the parser recognizes exactly the language described by the source grammar. This approach builds on the earlier work by Jacques-Henri Jourdan, François Pottier and Xavier Leroy on verified validation of $LR(1)$ parsers [60]. Jacques-Henri Jourdan succeeded in scaling this approach all the way up to the full ISO C99 grammar plus some extensions.
- Two new static analyses, value analysis and neededness analysis, were added to the CompCert back-end. As described in section 6.1.3 below, the results of these analyses enable more aggressive optimizations over the RTL intermediate form.
- As part of the work on formalizing floating-point arithmetic (see section 6.1.4 below), the semantics and compilation of floating-point arithmetic in CompCert was revised to handle single-precision floating-point numbers as first-class values, instead of systematically converting them to double precision before arithmetic. This increases the efficiency and compactness of the code generated for applications that make heavy use of single precision.
- Previously, the CompCert back-end compiler was assuming a partitioned register set from the target architecture, where integer registers always contain 32-bit integers or pointers, and floating-point registers always contain double-precision FP numbers. This convention on register uses simplified the verification of CompCert, but became untenable with the introduction of single-precision FP numbers as first-class values: FP registers can now hold either single- or double-precision FP numbers. Xavier Leroy rearchitected the register allocator and the stack materialization passes of CompCert, along with their soundness proofs, to lift this limitation on register uses. Besides mixtures of single- and double-precision FP numbers, this new architecture makes it possible to support future target processors with a unified register set, such as the SPE variant of PowerPC.
- We added support for several features of ISO C99 that were not handled previously: designated initializers, compound literals, `switch` statements where the `default` case is not the last case, `switch` statements over arguments of 64-bit integer type, and incomplete arrays as the last member of a `struct`. Also, variable-argument functions and the `<stdarg.h>` standard include are now optionally supported, but their implementation is neither specified nor verified.
- The ARM back-end was extended with support for the EABI-HF calling conventions (passing FP arguments and results in FP registers instead of integer registers) and with generation of Thumb2 instructions. Thumb2 is an alternate instruction set and instruction encoding for the ARM architecture that results in more compact machine code (up to 30% reduction in code size on our tests).

We released three versions of CompCert, integrating these enhancements: version 2.2 in February 2014, version 2.3 in April, and version 2.4 in September.

In June 2014, Inria signed a licence agreement with AbsInt Angewandte Informatik GmbH, a software publisher based in Saarbrucken, Germany, to market and provide support for the CompCert formally-verified C compiler. AbsInt will extend CompCert to improve its usability in the critical embedded software market, and also provide long-term maintenance as required in this market.

6.1.3. Value analysis and neededness analysis in CompCert

Participant: Xavier Leroy.

Xavier Leroy designed, implemented, and proved sound two new static analyses over the RTL intermediate representation of CompCert. Both analyses are of the intraprocedural dataflow kind.

- Value analysis is a forward analysis that tracks points-to information for pointers, constantness information for integer and FP numbers, and variation intervals for integer numbers, using intervals of the form $[0, 2^n)$ and $[-2^n, 2^n)$. This value analysis extends and generalizes CompCert's earlier

constant analysis as well as the points-to analysis of Robert and Leroy [68]. In particular, it tracks both the values of variables and the contents of memory locations, and it can take advantage of points-to information to show that function-local memory does not escape the scope of the function.

- Neededness analysis is a backward analysis that tracks which memory locations and which bits of the values of integer variables may be used later in a function, and which memory locations and integer bits are "dead", i.e. never used later. This analysis extends CompCert's earlier liveness analysis to memory locations and to individual bits of integer values.

Compared with the static analyses developed as part of Verasco (section 6.1.1), value analysis is much less precise: every function is analyzed independently of its call sites, relations between variables are not tracked, and even interval analysis is coarser (owing to CompCert's lack of support for widened fixpoint iteration). However, CompCert's static analyses are much cheaper than Verasco's, and scale well to large source codes, making it possible to perform them at every compilation run.

Xavier Leroy then modified CompCert's back-end optimizations to take advantage of the results of the two new static analyses, thus improving performance of the generated code:

- Common subexpression elimination (CSE) takes advantage of non-aliasing information provided by value analysis to eliminate redundant memory loads more aggressively.
- Many more integer casts (type conversions) and bit masking operations are discovered to be redundant and eliminated.
- Memory stores and block copy operations that become useless after constant propagation and CSE can now be eliminated entirely.

6.1.4. *Verified compilation of floating-point arithmetic*

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

In 2012, 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. This verification of FP arithmetic and of its compilation was further improved in 2013 with respect to the treatment of "Not a Number" special values.

This year, Guillaume Melquiond improved the algorithmic efficiency of some of the executable FP operations provided by Flocq. Xavier Leroy generalized the theorems over FP arithmetic used in CompCert's soundness proof so that these theorems apply both to single- and double-precision FP numbers. Jacques-Henri Jourdan and Xavier Leroy proved additional theorems concerning conversions between integers and FP numbers.

A journal paper describing this 3-year work on correct compilation of floating-point arithmetic was accepted for publication at Journal of Automated Reasoning [14].

6.1.5. *Verified JIT compilation of Coq*

Participants: Maxime Dénès, Xavier Leroy.

Evaluation of terms from Gallina, the functional language embedded within Coq, plays a crucial role in the performance of proof checking or execution of verified programs, and the trust one can put in them. Today, Coq provides various evaluation mechanisms, some internal, in the kernel, others external, via extraction to OCaml or Haskell. However, we believe that the specific performance trade-offs and the delicate issues of trust are still calling for a better, more adapted, treatment.

That is why we started in October this year the Coqonut project, whose objective is to develop and formally verify an efficient, compiled implementation of Coq reductions. As a first step, we wrote an unverified prototype in OCaml producing x86-64 machine code using a monadic intermediate form. We started to port it to Coq and to specify the semantics of the source, target and intermediate languages.

6.2. Language design and type systems

6.2.1. The Mezzo programming language

Participants: Thibaut Balabonski, François Pottier, Jonathan Protzenko.

Mezzo is a programming language proposal whose untyped foundation is very much like OCaml (i.e., it is equipped with higher-order functions, algebraic data structures, mutable state, and shared-memory concurrency) and whose type system offers flexible means of describing ownership policies and controlling side effects.

In 2013 and early 2014, Thibaut Balabonski and François Pottier re-worked the machine-checked proof of type soundness for Mezzo. They developed a version of the proof which includes concurrency and dynamically-allocated locks, and showed that well-typed programs do not crash and are data-race free. This work was presented by François Pottier at FLOPS 2014 [24]. The proof was then extended with a novel and simplified account of adoption and abandon, a mechanism for combining the static ownership discipline with runtime ownership tests. A comprehensive paper, which contains both a tutorial introduction to Mezzo and a description of its formal definition and proof, was submitted to TOPLAS.

Minor modifications were carried out by Jonathan Protzenko in the implementation. A version of Mezzo that runs in a Web browser was developed and uploaded online, so that curious readers can play with the language without installing the software locally.

Jonathan Protzenko wrote his Ph.D. dissertation [12], which describes the design of Mezzo and the implementation of the Mezzo type-checker. He defended on September 29, 2014.

Web site: http://protz.github.io/mezzo/

6.2.2. System F with coercion constraints

Participants: Julien Cretin [Trust In Soft], Didier Rémy, Gabriel Scherer.

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 study of such systems is to view type conversions as *coercions* inside terms.

Following a general approach based on System F, Julien Cretin and Didier Rémy earlier introduced a language of *explicit coercions* enabling abstraction over coercions and viewing all type transformations as explicit coercions [57]. To ensure that coercions are erasable, *i.e.*, that they only decorate terms without altering their reduction, they are restricted to those that are parametric in either their domain or codomain. Despite this restriction, this language already subsumed many extensions of System F, including bounded polymorphism, instance-bounded polymorphism, and η-conversions—but not subtyping constraints.

To lift this restriction, Julien Crétin and Didier Rémy proposed a new approach were coercions are left implicit. Technically, we extended System F with a rich language of propositions containing a first-order logic, a coinduction mechanism, consistency assertions, and coercions (which are thus just a particular form of propositions); we then introduced a type-level language using kinds to classify types, and constrained kinds to restrict kinds to types satisfying a proposition. Abstraction over types of a constrained kind amounts to abstraction over arbitrary propositions, including coercions.

By default, type abstraction should be erasable, which is the case when kinds of abstract type variables are inhabited—we say that such abstractions are consistent. Still, inconsistent type abstractions are also useful, for instance, to model GADTs. We provide them as a different construct, since they are not erasable, as they must delay reduction of subterms that depend on them. This introduces a form of weak reduction in a language with full reduction, which is a known source of difficulties: although the language remains sound, we loose the subject reduction property. This work has been described in [28] and is part of Julien Cretin's PhD dissertation [11] defended in January 2014; a simpler, core subset is also described in [45].

Recently, Gabriel Scherer and Didier Rémy introduced *assumption hiding* [32], [50] to restore confluence when mixing full and weak reductions and provide a continuum between consistent and inconsistent abstraction. Assumption hiding allows a fine-grained control of dependencies between computations and the logical hypotheses they depend on; although studied for a language of coercions, the solution is more general and should be applicable to any language with abstraction over propositions that are left implicit, either for the user's convenience in a surface language or because they have been erased prior to computation in an internal language.

6.2.3. Singleton types for code inference

Participants: Gabriel Scherer, Didier Rémy.

We continued working on 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. A decision algorithm for the simply-typed lambda-calculus is still work-in-progress. We presented at the TYPES'14 conference [42] our general approach to such decision procedures, and obtained an independent counting result for intuitionistic logic [52] that demonstrates the finiteness of the search space.

6.2.4. Generic programming with ornaments

Participants: Pierre-Évariste Dagand, Didier Rémy, Thomas Williams.

Since their first introduction in ML, datatypes have evolved: besides providing an organizing *structure* for computation, they are now offering more *control* over what is a valid result. GADTs, which are now part of the OCaml language, offer such a mechanism: ML programmers can express fine-grained, logical invariants of their datastructures. Programmers thus strive to express the correctness of their programs in the types: a well-typed program is correct by construction. However, these carefully crafted datatypes are a threat to any library design: the same data-*structure* is used for many logically incompatible purposes. To address this issue, McBride developed *ornaments*. It defines conditions under which a new datatype definition can be described as an ornament of another one, typically when they both share the same inductive definition scheme. For example, lists can be described as the ornament of the Church encoding of natural numbers. Once a close correspondence between a datatype and its ornament has been established, certain kinds of operations on the original datatype can be automatically lifted to its ornament.

To account for whole-program transformations, we developed a type-theoretic presentation of *functional ornament* [17] as a generalization of ornaments to functions. This work built up on a type-theoretic *universe of datatypes*, a first-class description of inductive types within the type theory itself. Such a presentation allowed us to analyze and compute over datatypes in a transparent manner. Upon this foundation, we formalized the concept of functional ornament by another type-theoretic universe construction. Based on this universe, we established the connection between a base function (such as addition and subtraction) and its ornamented version (such as, respectively, the concatenation of lists and the deletion of a prefix). We also provided support for driving the computer into semi-automatically lifting programs: we showed how addition over natural numbers could be incrementally evolved into concatenation of lists.

Besides the theoretical aspects, we have also tackled the practical question of offering ornaments in an ML setting [33]. Our goal was to extend core ML with support for ornaments so as to enable semi-automatic program transformation and fully-automatic code refactoring. We thus treated the purely syntactic aspects, providing a concrete syntax for describing ornaments of datatypes and specifying the lifting of functions. Such lifting specifications allow the user to declaratively instruct the system to, for example, lift addition of numbers to concatenation of lists. We gave an algorithm that, given a lifting specification, performs the actual program transformation from the bare types to the desired, ornamented types. This work has been evaluated by a prototype implementation in which we demonstrated a few typical use-cases for the semi-automatic lifting of programs.

Having demonstrated the benefits of ornaments in ML, it has been tempting to offer ornaments as first-class citizens in a programming language. Doing so, we wished to rationalize the lifting of programs as an elaboration process within a well-defined, formal system. To describe the liftings, one would like to specify only the local transformations that are applied to the original program. Designing such a language of *patches* and formalizing its elaboration has been the focus of our recent efforts.

6.2.5. Constraints as computations

Participant: François Pottier.

Hindley-Milner type inference–the problem of determining whether an ML program is well-typed–is well-understood, and can be elegantly explained and implemented in two phases, namely constraint generation and constraint solving. In contrast, elaboration–the task of constructing an explicitly-typed representation of the program–seems to have received relatively little attention in the literature, and did not until now enjoy a modular constraint-based presentation. François Pottier proposed such a presentation, which views constraints as computations and equips them with the structure of an applicative functor. This work was presented as a "functional pearl" at ICFP 2014 [31]. The code, in the form of a re-usable library, is available online.

6.2.6. Equivalence and normalization of lambda-terms with sums

Participants: Gabriel Scherer, Guillaume Munch-Maccagnoni [Université 13, LIPN lab].

Determining uniqueness of inhabitants requires a good understanding of program equivalence in presence of sum types. In yet-unpublished work, Gabriel Scherer worked on the correspondence between two existing normalization techniques, one coming from the focusing community [54] and the other using direct lambda-term rewriting [63]. A collaboration with Guillaume Munch-Maccagnoni has also started this year, whose purpose is to present normalization procedures for sums using System L, a rich, untyped syntax of terms (or abstract machines) for the sequent calculus.

6.2.7. Computational interpretation of realizability

Participants: Pierre-Évariste Dagand, Gabriel Scherer.

We are trying to better understand the computational behavior of semantic normalization techniques such as a realizability and logical relation models. As a very first step, we inspected the computational meaning of a normalization proof by realizability for the simply-typed lambda-calculus. It corresponds to an evaluation function; the evaluation order for each logical connective is determined by the definition of the sets of truth and value witnesses. This preliminary work is to be presented at JFLA 2015 [35].

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 2013 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 between 2010 and 2013. As part of this project, we are developing a C++ library, called PASL, for programming parallel computations at a high level of abstraction. We use this library to evaluate new algorithms and data structures. We obtained two major results this year.

The first result is a sequence data structure that provides amortized constant-time access at the two ends, and logarithmic time concatenation and splitting at arbitrary positions. These operations are essential for programming efficient computation in the fork-join model. Compared with prior work, this novel sequence data structure achieves excellent constant factors, allowing it to be used as a replacement for traditional, non-splittable sequence data structures. This data structure, called *chunked sequence* due to its use of chunks (fixed-capacity arrays), has been implemented both in C++ and in OCaml, and shown competitive with state-of-the art sequence data structures that do not support split and concatenation operations. This work is described in a paper published at ESA [22].

A second main result is the development of fast and robust parallel graph traversal algorithms, more precisely for parallel BFS and parallel DFS. The new algorithms leverage the aformentioned sequence data structure for representing the set of edges remaining to be visited. In particular, it uses the split operation for balancing the edges among the several processors involved in the computation. Compared with prior work, these new algorithms are designed to be efficient not just for particular classes of graphs, but for all input graphs. This work has not yet been published, however it is described in details in a technical report [46].

6.3.2. Weak memory models

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

Modern multi-core and multi-processor computers do not follow the intuitive "Sequential Consistency" model that would define a concurrent execution as the interleaving of the 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 instructions, 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 Microsoft Research, Cambridge), Peter Sewell (University of Cambridge) and Susmit Sarkar (University of St. Andrews) achieved several significant results, including two semantics for the IBM Power and ARM memory models: one of the operational kind [70] and the other of the axiomatic kind [64]. 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 2014 we produced 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 **herd** simulator (part of **diy** tool suite) is in fact a generic simulator, whose central component is an interpreter for a domain-specific language. More precisely, memory models are described in a simple language that defines relations by means of a few operators such as concatenation, transitive closure, fixpoint, etc., and performs validity checks on relations such as acyclicity. The Power/ARM model consists of about 50 lines of this specific language. This work, with additional material, including in-depth testing of ARM devices and data-mining of potential concurrency bugs in a huge code base, was published in the journal *Transaction on Programming Languages and Systems* [13] and selected for presentation at the PLDI conference [23]. Luc Maranget gave this presentation.

In the same research theme, Luc Maranget supervised the internship of Jacques-Pascal Deplaix (EPITECH), from Oct. 2013 to May 2014. Jacques-Pascal extended **litmus**, our tool to run tests on hardware. **litmus** now accepts test written in C; we can now perform the conformance testing of C compilers and machines with respect to the C11/C++11 standard. Namely, Mark Batty (University of Cambridge), under the supervision of Jade Alglave, wrote a **herd** model for this standard. The new **litmus** also proves useful to run tests that exploit some machine idiosyncrasies, when our **litmus** assembly implementation does not handle them.

As a part of the **litmus** infrastructure, Luc Maranget designed a synchronisation barrier primitive by simplifying the sense synchronisation barrier published by Maurice Herlily and Nir Shavit in their textbook [58]. He co-authored a JFLA article [34], that presents this primitive and proves it correct automatically by the means of the **cubicle** tool developed under the supervision of Sylvain Conchon (team Toccata, Inria Saclay).

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, Luc Maranget, Gabriel Scherer, Mark Shinwell [Jane Street], Leo White [OCaml Labs, Cambridge University], Jeremy Yallop [OCaml Labs, Cambridge University].

This year, we released versions 4.02.0 and 4.02.1 of the OCaml system. Release 4.02.0 is a major release that fixes about 60 bugs and introduces 70 new features suggested by users. Damien Doligez acted as release manager for both versions.

OCaml 4.02.0 introduces a large number of major innovations:

- Extension points: a uniform syntax for adding attributes and extensions in OCaml source code: most external preprocessors can now extend the language without need to extend the syntax and reimplement the parser.
- Improvements to the module system: generative functors and module aliases facilitate the efficient handling of large code bases.
- Separation between text-like read-only strings and array-like read-write byte sequences. This makes OCaml programs safer and clearer.
- An extension to the pattern-matching syntax to catch exceptions gives a short, readable way to write some important code patterns.
- Extensible open datatypes generalize the exception type and make its features available for general programming.
- Several important optimizations were added or enhanced: constant propagation, common subexpression elimination, dead code elimination, optimization of pattern-matching on strings.
- A code generator for the new 64-bit ARM architecture "AArch64".
- A safer and faster implementation of the printf function, based on the GADT feature introduced in OCaml 4.00.0.

This version has also seen a reduction in size: the Camlp4 and Labltk parts of the system are now independent systems. This makes them free to evolve on their own release schedules, and to widen their contributor communities beyond the core OCaml team.

OCaml 4.02.1 fixes a few bugs introduced in 4.02.0, along with 25 older bugs.

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

- Ephemerons: a more powerful version of weak pointers.
- A parallel extension of the runtime system and associated language features that will let multi-threaded OCaml programs run in parallel on several CPU cores.
- Modular implicits: a typeclass-like extension that will make is easy to write generic code (*e.g.* print functions, comparison predicates, overloaded arithmetic operators, etc).
- "Inlined" records as constructor arguments, which will let the programmer select a packed representation for important data structures.
- Major improvements to the inlining optimization pass.
- Support for debugging native-code OCaml programs with GDB.

6.4.2. Namespaces for OCaml

Participants: Fabrice Le Fessant, Pierrick Couderc.

With the growth of the OCaml community and the ease of sharing code through OPAM, the new OCaml package manager, OCaml projects are using more and more external libraries. As a consequence, conflicts between module names of different libraries are now more likely for big projects, and the need for switching from the current flat namespace to a hierarchical namespace is now real.

We experimented with a prototype of OCaml where the namespaces used by a module are explicitely written in the OCaml module source header, to generate the environment in which the source is typed and compiled [39]. Namespaces are mapped on directories on the disk. This mechanism complements the recent addition of module aliases to OCaml, by providing extensibility at the namespace level, whereas it is absent at the module level, and solves also the problem of exact dependency analysis (the previous tool used for that purpose, ocamldep, provides only an approximation of the dependencies, computed on the syntax tree).

6.4.3. Memory profiling OCaml application

Participants: Fabrice Le Fessant, Çagdas Bozman [ENSTA ParisTech], Grégoire Henry [OCamlPro], Michel Mauny [ENSTA ParisTech].

Most modern languages make use of automatic memory management to discharge the programmer from the burden of allocating and releasing the chunks of memory used by the software. As a consequence, when an application exhibits an unexpected usage of memory, programmers need new tools to understand what is happening and how to solve such an issue. In OCaml, the compact representation of values, with almost no runtime type information, makes the design of such tools more complex.

We have experimented with three tools to profile the memory usage of real OCaml applications. The first tool saves snapshots of the heap after every garbage collection. Snapshots can then be analysed to display the evolution of memory usage, with detailed information on the types of values, where they were allocated and from where they are still reachable. A second tool updates counters at every garbage collection event, it complements the first tool by providing insight on the behavior of the minor heap, and the values that are promoted or not to the major heap. Finally, a third tool samples allocations and saves stacks of function calls at these samples.

These tools have been used on real applications (Alt-Ergo, an SMT solver, or Cumulus, an Ocsigen website), and allowed us to track down and fix memory problems with these applications, such as useless copies of data structures and memory leaks.

6.4.4. OPAM, the OCaml package manager

Participants: Fabrice Le Fessant, Roberto Di Cosmo [IRILL], Louis Gesbert [OCamlPro].

With the growth of the OCaml community, the need for sharing libraries between users has lead to the development of a new package manager, called OPAM. OPAM is based on Dose, a library developed by the Mancoosi team at IRILL, to provide a unified format, CUDF, to query external dependency solvers. The specific needs of OPAM have driven interesting research and improvements on the Dose library, that have consequently opened new opportunities for improvements in OPAM, for the benefit of both software.

We have for example experimented with the design of a specific language [37] to describe optimization criteria, when managing OPAM packages. Indeed, depending on the actions (installation, upgrade, removal), the user might want to reach very different configurations, requiring an expressive power that go far beyond what traditional package managers can express in their configuration options. For example, during installation, the user would probably see as little compilation as possible, whereas upgrading is supposed to move the configuration to the most up-to-date state, with as much compilation as needed.

We have also proposed a new paradigm: multi-switch constraints, to model *switches* used in OPAM to handle different versions of OCaml on the same computer [41]. We proposed this new paradigm as a way to solve multiple problems (cross-compilation, multi-switch packages, per-switch repositories and application-specific switches). However, we expect this new paradigm to challenge the scalability of the current CUDF solvers used by OPAM, and to require important changes and optimization in the Dose library.

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 [61], and to build tools for writing TLA+ specifications and mechanically checking the corresponding formal proofs.

This year, we released two versions of the TLA+ Proof System (TLAPS), the part of the TLA+ tools that handles mechanical checking of TLA+ proofs. This environment is described in [55].

These versions add the propositional temporal logic prover LS4 as a back-end, which allows TLAPS to deal with propositional temporal formulas. This relies on a technique called *coalescing* [40], which allows users to prove arbitrary safety properties, as well as some liveness properties, by translating them into the back-end prover's logic without increasing the complexity of the formulas.

Jael Kriener started a post-doc contract in December 2013, funded by the ADN4SE contract, and left in September 2014. She worked on the theory of temporal proofs in TLA+ and, in collaboration with CEA, on proving some properties of the PharOS real-time operating system.

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 [Equipe DEDUCTEAM], Guillaume Bury [Equipe DEDUCTEAM], Olivier Hermant [Mines ParisTech].

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

Pierre Halmagrand continued his thesis work, funded by ANR BWare, on integrating Deduction Modulo in Zenon, with emphasis on making it efficient for dealing with B set theory.

Guillaume Bury did an internship, also funded by ANR BWare. He implemented an extension of Zenon, based on the simplex method, to deal with arithmetic formulas.

6.5.3. Well-typed generic fuzzing for module interfaces

Participants: Thomas Braibant, Jonathan Protzenko, Gabriel Scherer.

Property-based testing generates arbitrary instances of inputs to check a given correctness predicate/property. Thomas Braibant proposed that, instead of a random generation function defined from the internals of one's data-structure, one could use the user-exposed interface to generate instances by composition of API calls. GADTs let us reflect/reify a typed API, and program a type-respecting exploration/testing directly in the host language. We developed a prototype library, Articheck, to experiment with this idea. This work was presented at the ML Workshop [38].

6.5.4. Depth-First Search and Strong Connectivity in Coq

Participant: François Pottier.

In 2002, Ingo Wegener published a short paper which sketches a proof of Kosaraju's linear-time algorithm for computing the strongly connected components of a directed graph. At the same time, Wegener's paper helps explain why the algorithm works, which, from a pedagogical standpoint, makes it quite valuable. In 2013 and 2014, François Pottier produced a machine-checked version of Wegener's proof, and wrote a precise informal account of it, which will be presented at JFLA 2015 [36].

6.5.5. 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 makes 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.

We published an article in Journal of Automated Reasoning [15], explaining our work on this subject during the last 3 years. We gave four different approaches for implementing 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.

MUTANT Project-Team

6. New Results

6.1. Highlights of the Year

Acoustical Society of America Best Paper Award for [20].

International Computer Music Conference (ICMC) Best Presentation Award for [19].

MuTant TEDx Talk in October 2014 on *Human-Computer Musicianship* that attracted more than 12 thousand podcasts according to organisers.

MuTant in CNRS's 2nd edition of "Les Fondamentales" Science and Society event in Grenoble, in a session dedicated to Science and Music on the same Score.

MuTant Participation in the 2014 edition of *Futur en Seine* festival and showcased collaboration with Orchestre de Paris in a public event.

BEST PAPER AWARD :

[19] **International Computer Music Conference**. C. TRAPANI, J. ECHEVESTE.

6.2. Time-Coherency of Bayesian Priors for Sequential Alignment

In the context of Philippe Cuvillier's PhD project, we aim at increasing the robustness of machine listening in situations where observations from the external environment are extremely noisy or incoherent.

Recent results propose a novel insight to the problem of duration modeling for Information Retrieval problems where a discrete sequence of events is estimated from a time-signal using Bayesian models. Since the duration of each event is unknown, a major issue is setting the right Bayesian prior on each of them. Hidden Semi-Markov models (HSMM) allow choosing explicitly any probability distribution for the durations but learning these statistically is a non-parametric problem. In absence of huge training data sets, most algorithms rely on regularization techniques such as choosing parametric classes of distributions but the justifications of such techniques are often heuristics.

Among the numerous application domains of HMM-like paradigms, music-to-audio alignment brings two interesting properties. Firstly, a music score informs of the ordering among events. Secondly, it assigns to each event a nominal duration. For alignment tasks the Markov models conveniently model the ordering with *transient chains*. But the modeling of these nominal durations is a crucial and undermined problematic. This work investigates the relationship of this prior information of duration with the Bayesian priors of a HSMM. Theoretical insights are obtained through the study of the *prior state probability* of transient semi-Markov chains. Whereas ergodic chain and their convergence to an equilibrium probability are well studied, transient chains constitute an undermined case but of prime importance for real-time inference on HSMM.

On the first hand we prove that the non-asymptotical evolution of the state probability has some particular behaviors if the Bayesian priors fulfill several precise conditions, derived from statistical properties like the hazard rate and the tail decay. Then we say that a model is *time-coherent* if the evolution of the state probability respects the information of ordering and nominal lengths. This leads to several prescriptions on the design of HSMM Bayesian priors. On the other hand we get further prescriptions by comparing the Bayesian priors associated to different nominal lengths. This real-valued parameter comes with a natural ordering; we explain why this ordering among parameters is coherently modeled by some specific stochastic orderings among distributions that are standard in statistics.

Intermediate results have been reported in [12], [13]. This worked allowed the development of *Antescofo* version 0.6 released in November 2014.

6.3. Online Methods for Audio Segmentation and Clustering

Audio segmentation is an essential problem in many audio signal processing tasks, which tries to segment an audio signal into homogeneous chunks. Rather than separately finding change points and computing similarities between segments, we focus on joint segmentation and clustering, using the framework of hidden Markov and semi-Markov models. We introduced a new incremental EM algorithm for hidden Markov models (HMMs) and showed that it compares favorably to existing online EM algorithms for HMMs. Early experimental results on musical note segmentation and environmental sound clustering are promising and will be pursued further in 2015.

This project was done in the context of Alberto Bietti's MS project [26] under co-supervision of Arshia Cont (MuTant) and Francis Bach (SIERRA).

6.4. Model-based Testing an Interactive Music System

In the context of the Phd of Clément Poncelet, and in relation with the developments presented in Section 5.3 , we have been studying the application of model-based timed testing techniques to interactive music systems like Antescofo.

Several formal methods have been developed for automatic conformance testing of critical embedded software. The principle is to execute a real implementation under test (IUT) in a testing framework, black-box, by sending it carefully selected inputs and then observing and analyzing its outputs. In conformance model-based testing (MBT), the input and corresponding expected outputs are generated according to formal models of the IUT and the environment. The models of timed automata with inputs and outputs, and tools like the the Uppaal suite have been developed for extending such techniques to realtime systems [32], [31]. Several procedures have been designed for addressing the task described in Section 5.3 .

The case of IMS presents important originalities compared to other MBT applications to realtime systems. On the one hand, the time model supports several time units, including the wall clock time, measured in seconds, and the time of music scores, measured in number of beats relatively to a tempo. This situation raised several new problems for the generation of test suites and their execution. On the other hand, the formal specification of the IUT's behavior on a given score is produced automatically by a *score compiler*, using an intermediate representation. We rely on the realistic hypotheses that a mixed score specify completely the expected timed behavior of the IMS. Hence, our test method is fully automatic, in contrast with other approaches which generally require experts to write the specification manually. This workflow fits well in a music authoring workflow where scores in preparation are constantly evolving. We have been applying our tools to small benchmark made of characteristic scores, as well as to real mixed scores used in concerts, and some bugs in Antescofo have been identified. These results have been presented in the conference ICMC 2014 [18] and will be presented during the 30th ACM/SIGAPP Symposium On Applied Computing, track Software Verification and Testing [17].

6.5. Antescofo Temporal Pattern

An important enhancement has been made by the introduction of an expressive temporal pattern language [15] in *Antescofo*. Temporal patterns are used to define complex events that correspond to a combination of perceived events in the musical environment as well as arbitrary logical and metrical temporal conditions. The real time recognition of such event is used to trigger arbitrary actions in the style of event-condition-action rules.

The semantics of temporal pattern matching is defined to parallel the well-known notion of regular expression and Brzozowski's derivatives but extended to handle an infinite alphabet, arbitrary predicates, elapsing time and inhibitory conditions.

Temporal patterns are implemented by translation into a core subset of the Antescofo domain specific language. This compilation has proven efficient enough to avoid the extension of the real-time runtime of the language and has been validated with composers in actual pieces.

6.6. OpenMusic reactive Model

In collaboration with Jean Bresson, we have extended the evaluation model of OpenMusic to integrate reactive capabilities [10]. OpenMusic (OM) is a domain-specific visual programming language designed for computer-aided music composition based on Common Lisp. It allows composers to develop functional processes generating or transforming musical data. To extend OM towards reactive applications, we have proposed to integrate its demand-driven evaluation mechanism with reactive data-driven evaluations in a same and consistent visual programming framework. To this end, we have developped the first denotational semantics of the visual language, which gives account for its demand-driven evaluation mechanism and the incremental construction of programs. We then have extended this semantics to enable reactive computations in the functional graphs. The resulting language merges data-driven executions with the existing demand-driven mechanism. This integration allows for the propagation of changes in the programs, and the evaluation of graphically-designed functional expressions as a response to external events, a first step in bridging the gap between computer-assisted composition environments and real-time musical systems.

6.7. Representation of Rhythm and Quantization

Rhythmic data are commonly represented by tree structures (rhythms trees) in assisted music composition environments, such as OpenMusic, due to the theoretical proximity of such structures with traditional musical notation. We are studying the application in this context of techniques and tools for processing tree structure, which were originally developed for other areas such as natural language processing, automatic deduction, Web data ... We are particularly interested in two well established formalisms with solid theoretical foundations: term rewriting and tree automata.

The problem of rhythm transcription, or quantization, is to generate, from a timed sequence of notes (e.g. a file in MIDI format), a score in traditional music notation. The input events can come from an interpretation on a MIDI keyboard or be the result of a computation in OpenMusic. This problem arises immediately as insoluble unequivocally: we shall calibrate the system to fit the musical context, balancing constraints of precision, or of simplicity / readability of the generated scores. For this purpose, we are developing in collaboration with Slawek Staworko (LINKS, currently on leave at University of Edinburgh) for algorithms searching optimums in large sets of weighted trees (tree series), representing possible solutions to a problem quantification. A prototype has been developed and is under evaluation on real case studies. For the construction of appropriate tree series, we turn to semi-supervised systems, where the composer's interactions are predominant in the smooth process. These work have been presented in an invited talk in the workshop of the IFIP working group on term rewriting.

With Prof. Masahiko Sakai (Nagoya University, a specialist in term rewriting), we conduct a complementary work [14] on the representation of rhythmic notation. The goal is to define a structural theory as equations on trees rhythms. This approach can be used for example to generate, by transformation, different notations possible the same rate, with the ability to select in accordance with certain constraints.

PARKAS Project-Team

6. New Results

6.1. Highlights of the Year

The paper *ReactiveML, a reactive extension to ML* of Mandel and Pouzet has been declared to be the *most influential paper of PPDP (Principles and Practice of Declaractive Programming) 2005*. A previous version of the paper, submitted to JFLA'05, has been declared to be "une contribution marquante parmi les articles publiés aux JFLA".

6.2. Quasi-synchrony

Participants: Guillaume Baudart, Timothy Bourke, Marc Pouzet.

We study the implementation of critical control applications on the so-called *quasi-periodic* distributed architectures. These architectures, used in civil avionics (e.g., Airbus A380), consist of a collection of distributed processors running with *quasi-periodic* clocks, that is, un-synchronized physical clocks subject to bounded jitterring. The theory of quasi-synchrony has been introduced by Paul Caspi in the 2000' [29]. Loosely Time-Triggered Architectures (LTTA) denotes such architectures with the prototocol used to implement a synchronous program on top of it.

Over the last ten year two protocols were considered: (1) *Back-Pressure* LTTA [25] based on a acknowledgement mechanism reminiscent of elastic circuit [43]. (2) *Time-Based* LTTA [28] which uses timing constraints of the architecture to mimic a synchronous execution.

During year 2014, we have entirelly reformulated the model of LTTA using synchronous semantics and principles. Compared to previous formalizations based on Petri nets [24], this new presentation is is simpler and more uniform with the same theoretical model used for both the application and the protocol ((1) or (2)). Moreover, it is easier to consider mixed protocols (a whole application with part based on time-based communication and others based on back-pressure). Besides this, we also proposed a new and more flexible Time-Based LTTA, allowing for pipelining by not reconstructing global synchronization, unlike what was done in previous Time-Based LTTA.

6.3. Hybrid Synchronous Languages

Participants: Guillaume Baudart, Timothy Bourke, Marc Pouzet.

During year 2014, we mainly worked on two directions: (a) the design and implementation of causality analysis for hybrid systems modelers; (b) the design and implementation of a new compilation technique producing imperative sequential code.

This research is conducted in collaboration with Albert Benveniste and Benoit Caillaud (Hycomes team at Inria, Rennes), Jean-Louis Colaco, Cédric Pasteur and Bruno Pagano from the SCADE core team of Esterel-Technologies/ANSYS.

Causality analysis In this work, we address the static detection of causality loops for a hybrid modeling language that combines synchronous Lustre-like data-flow equations with Ordinary Differential Equations (ODEs). We introduce the operator *last(x)* for the left-limit of a signal x. This operator is used to break causality loops and permits a uniform treatment of discrete and continuous state variables. The semantics relies on non-standard analysis, defining an execution as a sequence of infinitesimally small steps. A signal is deemed *causally correct* when it can be computed sequentially and only progresses by infinitesimal steps outside of discrete events. The causality analysis takes the form of a simple type system. In well-typed programs, signals are proved continuous during integration.

This analysis has been presented at [4] and is fully implemented in the hybrid synchronous language Zélus.

A Synchronous-based Code Generator For Explicit Hybrid Systems Languages The generation of sequential code is important for simulations to be efficient and to produce target embedded code. While sequential code generation in hybrid modeling tools is routinely used for efficient simulation, it is little or not used for producing target embedded code in critical applications submitted to strong safety requirements. This is a break in the development chain: parts of the applications must be rewritten into either sequential or synchronous programs, and all properties verified on the source model cannot be trusted and have to be re-verified on the target code.

In this work, we present a novel approach for the code generation of a hybrid systems modeling language. By building on top of an existing synchronous language and compiler, it reuses almost all the existing infrastructure with only a few modifications. Starting from an existing synchronous data-flow language extended with Ordinary Differential Equations (ODEs), we detail the translation to sequential code. The translation is expressed as a sequence of source-to-source transformations. A generic intermediate language is introduced to represent transition functions which are turned into C code. The versatility of the compiler organisation is illustrated by considering two classical targets: generation of simulation code complying with the FMI standard and linking with an off-the-shelf numerical solver (Sundials CVODE).

This new code generation has been implemented in two different compilers: the Zélus research prototype and the industrial SCADE Suite KCG code generator, at Esterel-Technologies/ANSYS. Here, SCADE is conservatively extended with ODEs, following previous works by Benveniste et al. and implemented in Zélus. In the SCADE compiler, it was possible to reuse almost all the existing infrastructure like static checking, intermediate languages, and optimisations, with few modifications. The extension to account for hybrid features represents only 5% additional lines of code, which is surprisingly low. Moreover, the proposed language extension is conservative in that regular synchronous functions are compiled as before—the same synchronous code is used both for simulation and for execution on target platforms.

This full-scale validation confirm the interest in building a hybrid systems modeler on top of a existing synchronous language. Moreover, the precise definition of code generation, built on a proven compiler infrastructure of a synchronous language avoids the rewriting of control software and may also increase the confidence in what is simulated.

This work will be presented at the *International Conference on Compiler Construction (CC)*, in April 2015.

6.4. Fidelity in Real-Time Programming

Participants: Guillaume Baudart, Timothy Bourke.

Synchronous languages are a rigorous approach to programming, analyzing, and implementing embedded systems. Real-time aspects are typically handled by discretizing time using either (implicit) ticks or (explicit) named signals, and later verifying that the (necessarily bounded) execution time of a reaction is strictly less than the period of the fastest timing signal. This approach has many advantages: it separates logical behaviour from implementation concerns, yields a simple and precise programming model, and abstracts from eventual run-time environments. For an important subclass of embedded protocols and controllers, however, we believe it advantageous to add constructions that deal more concretely with real-time constraints.

We are pursuing these ideas in the enriched timing model provided by the Zélus programming language (detailed elsewhere). We continue to study the extension and application of this language to the modelling, simulation, analysis, and implementation of real-time embedded software.

This year we developed three case studies: quasi-synchronous architectures (from last year), loosely time-triggered architectures (detailed elsewhere), and a small embedded controller. These case studies motivate and drive our research and implicitly define the subclass of embedded systems that we aim to treat. They have each been modelled in Zélus and can be simulated with the existing compiler.

We made progress on defining a subset of Zélus that is ammenable to discretization techniques for more flexible simulation. A first version of an appropriate algorithm has been sketched and partially implemented. Work continues on developing it with the idea of incorporating it into the Zélus compiler and using it to treat our case studies.

6.5. 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).

This year I finalized both the framework for network invariant proofs [20] and its application to the AODV protocol [21] and submitted them for inclusion in the *Archive of Formal Proof*, an online and open-source repository of formal developements in the Isabelle proof assistant (indexed as a journal). I presented results on the framework at the Vienna 'Summer of Logic' [6] and my colleagues presented the application in Sydney [5]. Together with an intern at NICTA and Sydney, my colleagues and I made preliminary investigations into extending the framework and model with timing details. A journal version of the ITP paper has been submitted.

In collaboration with Peter Höfner (NICTA) and Robert J. van Glabbeek (UNSW/NICTA).

6.6. Reasoning about C11 Program Transformations

Participants: Francesco Zappa Nardelli, Thibaut Balabonski, Robin Morisset.

We have shown that the weak memory model introduced by the 2011 C and C++ standards does not permit many of common source-to-source program transformations (such as expression linearisation and "roach motel" reordering) that modern compilers perform and that are deemed to be correct. As such it cannot be used to define the semantics of intermediate languages of compilers, as, for instance, LLVM aimed to. We consider a number of possible local fixes, some strengthening and some weakening the model. We have evaluated the proposed fixes by determining which program transformations are valid with respect to each of the patched models. We have provided formal Coq proofs of their correctness or counterexamples as appropriate.

A paper on this work has been accepted in [13]. In collaboration with Viktor Vafeiadis (MPI-SWS, Germany).

6.7. Language design on top of JavaScript

Participant: Francesco Zappa Nardelli.

This research project aims at improving the design of the JavaScript language. In [22] 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 have explored the performance and software engineering benefits.

With Gregor Richards and Jan Vitek (Purdue University).

6.8. Tiling for Stencils

Participants: Tobias Grosser, Sven Verdoolaege, Albert Cohen.

This research project aims with optimizing time-iterated stencil operations.

Iterative stencil computations are important in scientific computing and more and more also in the embedded and mobile domain. Recent publications have shown that tiling schemes that ensure concurrent start provide efficient ways to execute these kernels. Diamond tiling and hybrid-hexagonal tiling are two tiling schemes that enable concurrent start. Both have different advantages: diamond tiling has been integrated in a general purpose optimization framework and uses a cost function to choose among tiling hyperplanes, whereas the greater flexibility with tile sizes for hybrid-hexagonal tiling has been exploited for effective generation of GPU code.

We undertook a comparative study of these two tiling approaches and proposed a hybrid approach that combines them. We analyzed the effects of tile size and wavefront choices on tile-level parallelism, and formulate constraints for optimal diamond tile shapes. We then extended, for the case of two dimensions, the diamond tiling formulation into a hexagonal tiling one, which offers both the flexibility of hexagonal tiling and the generality of the original diamond tiling implementation. We also showed how to compute tile sizes that maximize the compute-to-communication ratio, and apply this result to compare the best achievable ratio and the associated synchronization overhead for diamond and hexagonal tiling.

One particularly exciting result is the ability to apply tiling to periodic data domains. These computations are prevalent in computational sciences, particularly in partial differential equation solvers. We proposed a fully automatic technique suitable for implementation in a compiler or in a domain-specific code generator for such computations. Dependence patterns on periodic data domains prevent existing algorithms from finding tiling opportunities. Our approach augments a state-of-the-art parallelization and locality-enhancing algorithm from the polyhedral framework to allow time-tiling of stencil computations on periodic domains. Experimental results on the swim SPEC CPU2000fp benchmark show a speedup of $5\times$ and $4.2\times$ over the highest SPEC performance achieved by native compilers on Intel Xeon and AMD Opteron multicore SMP systems, respectively. On other representative stencil computations, our scheme provides performance similar to that achieved with no periodicity, and a very high speedup is obtained over the native compiler. We also report a mean speedup of about $1.5\times$ over a domain-specific stencil compiler supporting limited cases of periodic boundary conditions. To the best of our knowledge, it has been infeasible to manually reproduce such optimizations on swim or any other periodic stencil, especially on a data grid of two-dimensions or higher.

These works resulted in a number of high-profile publications, including a nommination for a best paper award, and culminated with the PhD thesis defense of Tobias Grosser.

6.9. Portable representation for polyhedral compilation

Participants: Riyadh Baghdadi, Michael Kruse, Chandan Reddy, Tobias Grosser, Sven Verdoolaege, Albert Cohen.

Programming accelerators such as GPUs with low-level APIs and languages such as OpenCL and CUDA is difficult, error prone, and not performance-portable. Automatic parallelization and domain specific languages have been proposed to hide this complexity and to regain some performance portability. We proposed PENCIL, a subset of GNU C99 with specific programming rules. A compiler for a Domain-Specific Language (DSL) may use it as a target language, a domain expert may use it as a portable implementation language facilitating the parallelization of real-world applications, and an optimization expert may use PENCIL to accelerate legacy applications.

The design of PENCIL is simultaneously a key research result and a milestone for parallelizing compiler engineering/design. Aspects of its static-analysis-friendly, formal semantics are highly original, for the language's ability to preserve expressiveness and modularity without jeopardizing a (polyhedral) compiler's ability to perform aggressive transformations. We validated its potential as a front-end to a state-of-the-art polyhedral compiler, extending its applicability to dynamic, data dependent control flow and non-affine array accesses. We illustrated this PENCIL-enabled flow on the generation of highly optimized OpenCL code, considering a set of standard benchmarks (Rodinia and SHOC), image processing kernels, and DSL embedding scenarios for linear algebra (BLAS) and signal processing radar applications (SPEAR-DE). We ran experimental results on a variety of platforms, including an AMD Radeon HD 5670 GPU, an Nvidia GTX470 GPU, and an ARM Mali-T604 GPU.

This work is conducted in collaboration with partners from ARM, RealEyes (a computer vision company) and Imperial College.

6.10. Correct and efficient runtime systems

Participants: Nhat Minh Lê, Robin Morisset, Adrien Guatto, Albert Cohen.

Complementing our different compilation efforts for synchronous and task-parallel data-flow languages, we studied the implementation of Kahn process networks, a deterministic parallel programming model, on shared memory multiprocessors. This model is based on a familiar abstraction: blocking communication through bounded, in-order, single-producer single-consumer queues.

We proposed two novel algorithms that construct such blocking queues on top of concurrent ring buffers and user-land scheduling components. We implemented our algorithms in C11, taking advantage of the relaxed memory model of the language, and prove the correctness of this implementation.

We used these algorithms in a complete runtime system for Kahn process networks with applications ranging from linear algebra kernels to stream computing. In particular, our implementations of the Cholesky and LU factorizations outperform state-of-the-art parallel linear algebra libraries on commodity x86 hardware.

6.11. A Functional Synchronous Language with Integer Clocks

Participants: Adrien Guatto, Albert Cohen, Louis Mandel, Marc Pouzet.

Synchronous languages in the vein of Lustre are first-order functional languages dedicated to stream processing. Lustre compilers use a type-like static analysis, the clock calculus, to reject programs that cannot be implemented as finite state machines. The broad idea is to assign to each element of a stream a logical computation date in a global, discrete time scale. When this analysis succeeds, the types obtained guide the code generation phase of the compiler, which produces transition functions. In practice, these functions consists in simple, bounded memory C code featuring only assignments and conditional statements.

This research work explores a variation on Lustre and its compilation. Our proposal is twofold. First, we add a new construct that creates a local time scale whose internal steps are invisible from the outside. Second, we change the clock calculus to allow several elements of a stream to be computed during the same time step. The resulting type system comes with a soundness proof, which relies on an elementary form of step-indexed realizability, and with a code generation scheme adapted to the new setting, and featuring nested loops in the target code.

PI.R2 Project-Team

5. New Results

5.1. Highlights of the Year

We successfully organised the thematic trimester Semantics of Proofs and Certified Mathematics (IHP, April-July 2014). The trimester attracted over two hundred participants altogether (with about 60 "resident" participants staying a month or more), hosted 5 special workshops, as well as other related regevents such as Types, MAP (Mathematics, Algorithms, and Proofs). It was the first thematic trimester in the history of IHP to feature computer science prominently. There was a kick-off day on April 22, with talks of Georges Gonthier, Thomas Hales, Xavier Leroy, and Vladimir Voevodsky, with the presence of some science journalists. During the trimester, the Bourbaki Seminar devoted an afternoon (June 21) to these themes, with talks of Thomas Hales and Thierry Coquand.

Shortly before, Coq has received the Software System Award 2013 from the Association for Computing Machinery (ACM). Hugo Herbelin is one of the recipients of this prize.

5.2. Proof-theoretical and effectful investigations

Participants: Pierre Boutillier, Guillaume Claret, Pierre-Louis Curien, Amina Doumane, Hugo Herbelin, Etienne Miquey, Ludovic Patey, Pierre-Marie Pédrot, Yann Régis-Gianas, Alexis Saurin.

5.2.1. Proving with side-effects

In 2012, Hugo Herbelin showed that classical arithmetic in finite types extended with strong elimination of existential quantification proves the axiom of dependent choice. To get classical logic and choice together without being inconsistent is made possible first by constraining strong elimination of existential quantification to proofs that are essentially intuitionistic and secondly by turning countable universal quantification into an infinite conjunction of classical proofs evaluated along a call-by-need evaluation strategy so as to extract from them intuitionistic contents that complies to the intuitionistic constraint put on strong elimination of existential quantification. Étienne Miquey is currently working to get a presentation of this work in Curien-Herbelin's μ-$\widetilde{\mu}$-calculus, with the aim of getting in the end a CPS-translation. Such a translation would provide a strong argument of normalisation for the calculus, as well as a better undertanding of the mechanisms of the calculus, especially the side-effect part and the meaning of the existential quantifier restriction.

Hugo Herbelin and Danko Ilik carried on their work on the computational content of completeness proofs and in particular of the computational content of Gödel's completeness theorem. Hugo Herbelin presented their work at the workshop PSC 2014.

5.2.2. Reverse mathematics

Ludovic Patey studied with Laurent Bienvenu and Paul Shafer the provability strength of Ramsey-type versions of theorems like König's lemma. The corresponding paper is submitted to the Journal of Mathematical Logic. Ludovic Patey studied with Laurent Bienvenu the constructions of diagonal non-computable functions by probabilistic means. They submitted a paper to Information and Computation. Ludovic Patey worked on the existence of universal instances in reverse mathematics, and submitted a paper to Annals of Pure and Applied Logic. He worked on the relations between diagonal non-computability and Ramsey-type theorems and submitted a paper to the Archive for Mathematical Logic. He studied the links between the iterative forcing framework developed by Lerman, Solomon & Towsner and the notion of preservation of hyperimmunity and submitted a paper to Computability in Europe 2015.

5.2.3. Gödel's functional interpretation

Pierre-Marie Pédrot kept developing the proof-as-program interpretation of Gödel's Dialectica translation, as seen through the prism of classical realisability. This work was presented at TYPES 2014 and later published at LICS 2014 [26].

5.2.4. 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. This work was presented at JFLA 2014 [30] early 2014 and later expanded to the classical case, encompassing $\lambda\mu$-calculus.

5.2.5. 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. Their work gave rise to a presentation at the Workshop on Infinitary Rewriting that took place in Vienna last July as part of FLOC 2014.

5.2.6. Alternative syntaxes for proofs

Amina Doumane and Alexis Saurin, in a joint work with Marc Bagnol, studied the structure of several correctness criteria for linear logic proof-nets and could relate them through a new primitive notion of dependency. This work was first presented at JFLA 2014 [29] early 2014 and later at Structure and Deduction in Vienna as part of FLOC 2014. An expanded version has recently been accepted at FOSSACS 2015 [19].

5.3. Type theory and the foundations of Coq

Participants: Pierre Boutillier, Pierre-Louis Curien, Hugo Herbelin, Pierre-Marie Pédrot, Yann Régis-Gianas, Matthieu Sozeau, Arnaud Spiwack.

5.3.1. Description of type theory

Hugo Herbelin and Arnaud Spiwack completed and published their characterisation of the type constructions of Coq in terms of atomic constructions rather than their usual description as a monolithic scheme [23]. This work permitted both a more pedagogical presentation of Coq's type system, and a more tractable and composable mathematical model of Coq on which meta-properties can be stated and proved.

5.3.2. 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 is in the process of being published in a special issue of MSCS on homotopy type theory [9].

The technique scales to provide type-theoretic constructions for arbitrary presheaves on Reedy categories, thus including simplicial sets.

5.3.3. Proof irrelevance, eta-rules

During his master's internship supervised by Matthieu Sozeau, Philipp Haselwarter studied a formulation of proof-irrelevance based on the rooster and the syntactic bracket presentation by Spiwack and Herbelin [23]. This resulted in a decomposition of the calculus cleanly showing the use of smashing and a better understanding of the restricted elimination rules of propositions. It also clearly shows that the inductive type for accessibility, used to justify general wellfounded definitions, can not be interpreted as a proof-irrelevant proposition in this calculus.

5.3.4. Unification

Matthieu Sozeau is continuing work in collaboration with Beta Ziliani (PhD at MPI-Saarbrücken) on formalising the unification algorithm used in Coq, which is central for working with advanced type inference features like Canonical Structures. This is the first precise formalisation of all the rules of unification including the ones used for canonical structure resolution. The presentation currently excludes some heuristics that were added on top of the core algorithm in Coq, until they can be studied more carefully. This work, part of B. Ziliani's thesis, was presented at the UNIF'14 workshop [28] and the Coq workshop in Vienna. A submission is in preparation.

5.3.5. Foundations and paradoxes

Arnaud Spiwack generalised previous works by Herman Geuvers and Hugo Herbelin to implement Hurkens's paradox of the impredicative system U^-. The resulting Coq implementation, which is completely independent from the impredicative features of Coq, generalises the two special cases which were previously used to prove negative results about impredicativity in Coq.

5.4. Homotopy of rewriting systems

Participants: Cyrille Chenavier, Pierre-Louis Curien, Yves Guiraud, Maxime Lucas, Philippe Malbos, Jovana Obradović.

5.4.1. Coherent presentations of Artin monoids

With Stéphane Gaussent (ICJ, Univ. de Saint-Étienne), Yves Guiraud and Philippe Malbos have used higher-dimensional rewriting methods for the study of Artin monoids, a class of monoids that is fundamental in algebra and geometry. This work uses the formal setting of coherent presentations (a truncation of polygraphic resolutions at the level above relations) to formulate, in a common language, several known results in combinatorial group theory: one by Tits about the fundamental group of a graph associated to an Artin monoid [65], and one by Deligne about the actions of Artin monoids on categories [47], both proved by geometrical methods. In this work, an improvement of Knuth-Bendix's completion procedure is introduced, called the homotopical completion-reduction procedure, and it is used to give a constructive proof and to extend both theorems. This work will appear in Compositio Mathematica [18] and has been implemented in a Python library.

The next objective of this collaboration is to extend those results in every dimension, first to Artin monoids, then to Artin groups, with a view towards two well-known open problems in the field: the word problem of Artin groups and the so-called $K(\pi, 1)$ conjecture.

5.4.2. New methods for the computation of polygraphic resolutions

Maxime Lucas, supervised by Pierre-Louis Curien and Yves Guiraud, develops Squier's theory in the setting of cubical ω-categories. This will allow easier and more explicit computations of polygraphic resolutions than in the globular setting of [5], and the use of new effective methods such as the reversing algorithm from Garside theory [46].

Yves Guiraud currently collaborates with Patrick Dehornoy (Univ. de Caen) and Matthieu Picantin (LIAFA, Univ. Paris 7) to extend the constructions of [18] to other important families of monoids, such as the plactic monoid, the Chinese monoid and the dual braid monoids.

5.4.3. Higher-dimensional linear rewriting

Cyrille Chenavier, Pierre-Louis Curien, Yves Guiraud and Philippe Malbos investigate with Eric Hoffbeck (LAGA, Univ. Paris 13) and Samuel Mimram (LIX, École Polytechnique) the links between set-theoretic rewriting theory and the computational methods known in symbolic algebra, such as Gröbner bases [39]. This interaction is supported by the Focal project of the IDEX Sorbonne Paris Cité.

With Eric Hoffbeck (LAGA, Univ. Paris 13), Yves Guiraud and Philippe Malbos have introduced the setting of linear polygraphs to formalise a theory of linear rewriting (in the sense of linear algebra), generalising Gröbner bases. They have adapted to algebras the procedure of [5] that computes polygraphic resolutions from convergent presentations of monoids, with applications to the decision of an important homological property called Koszulness. This work is contained in [35] and it has been presented at IWC 2014 [31].

Cyrille Chenavier, supervised by Yves Guiraud and Philippe Malbos, explores the use of Berger's theory of reduction operators [38] to design new methods for the study of linear rewriting systems, and to promote the use of rewriting techniques in combinatorial algebra.

5.4.4. Homotopical and homological finiteness conditions

Yves Guiraud and Philippe Malbos have written a comprehensive introduction [36] 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.

5.4.5. Wiring structure of operads and operad-like structures

Building on recent ideas of Marcelo Fiore on the one hand, and of François Lamarche on the other hand, Pierre-Louis Curien and Jovana Obradović develop a syntactic approach, using some of the kit of Curien-Herbelin's duality of computation and its polarised versions of Munch and Curien, to the definition of various structures that have appeared in algebra under the names of operads, cyclic operads, dioperads, properads, modular and wheeled operads, permutads, etc.... These structures are defined in the literature in different flavours. We seek to formalise the proofs of equivalence between these different styles of definition, and to make these proofs modular, so as not to repeat them for each variation of the notion of operad. Preliminary results are being presented in January 2015 at the Mathematical Institute of the Academy of Sciences (Belgrade).

5.5. Coq as a functional programming language

Participants: Pierre Boutillier, Guillaume Claret, Lourdes Del Carmen González Huesca, Thibaut Girka, Hugo Herbelin, Pierre Letouzey, Matthias Puech, Yann Régis-Gianas, Matthieu Sozeau, Arnaud Spiwack.

5.5.1. Type classes and libraries

Type Classes are heavily used in the HoTT/Coq library (http://github.com/HoTT/coq) started by the Univalent Foundations program at the IAS, to which Matthieu Sozeau participated. To ease the development of this sophisticated library, Matthieu Sozeau implemented a number of extensions to type class resolution to make it more predictable and efficient. These are now part of the Coq 8.5 release.

5.5.2. Dependent pattern-matching

The dissertation of Pierre Boutillier presents and formalises a new algorithm to compile dependent pattern-matching into a chain of Coq case analyses. It avoids the use of the "uniqueness of identity proofs" axiom in more cases than the former proposal by McBride and McKinna.

5.5.3. Incrementality in proof languages

Lourdes del Carmen González Huesca and Yann Régis-Gianas developed a new variant of the differential lambda calculus that has two main features: (i) it is deterministic ; (ii) it is based on a notion of a first-class changes. A paper is in preparation.

5.5.4. Proofs of programs in Coq

In collaboration with David Mentre (Mitsubishi Rennes), Thibaut Girka and Yann Régis-Gianas worked on a certified generator for correlating programs. A correlating program is a program that represents the semantic difference between two (close) versions of a program by performing a static scheduling of their instructions. Performing an abstract interpretation on the correlating program provides a representation of the semantic differences between the two versions of a program. A paper is written and should be submitted soon.

5.5.5. Typed tactic language

In collaboration with Beta Ziliani (MPI) and Thomas Refis (master 2 student at University Paris Diderot), Yann Régis-Gianas starts the development of the version 2 of Mtac, a tactic language for Coq. Mtac is a DSL embedded in the Coq proof assistant. Roughly speaking, it allows Coq to be used as a tactic language for itself. With this work, Mtac 2 now includes first class goals. A paper is in preparation.

5.5.6. Tactic engine

Arnaud Spiwack joined the team for two months (Sept—Oct 2014) to finalise the integration and documentation of his re-engineering of Coq's interactive proof engine for the v8.5 version. The new perspective taken by this new engine is to shift the primary focus from how tactics (proof instructions) can modify goals (proof obligations) to focus on the way tactics compose. By making sure that composition of tactics has good mathematical properties, the new engine makes it possible to combine tactics in a more predictable and more powerful way. This new engine is also notable for the introduction of an abstract interface for tactics and tactic composition which makes it easy to augment tactics with new capabilities. The most notable such features are so-called dependent subgoals, which makes more fine-grained proofs possible and significantly improves the support for dependent types; and backtracking which gives the possibility to deploy very modular proof-search components. During his two months in the team, Arnaud Spiwack also added support for tracing tactic execution (Info), again taking advantage of his modular design.

5.5.7. Effectful programming

Guillaume Claret and Yann Régis-Gianas developed a compiler from a subset of OCaml with effects to Coq. Possible effects are the exceptions, the global references and the non-termination. Guillaume Claret and Yann Régis-Gianas developed Pluto, a concurrent HTTP web server written in Gallina. They worked on techniques to certify such interactive programs, formalising the reasoning by use cases. Use cases are proven correct giving a scenario, a typed schema of interactions between a program and an environment, built using the tactic mode of Coq as a symbolic debugger.

5.5.8. Libraries

Sébastien Hinderer and Pierre Letouzey contributed an extended library of lists. Pierre Letouzey contributed an extended library about Peano numbers, that takes advantages of the "Numbers" modular framework done earlier.

POLSYS Project-Team

6. New Results

6.1. Highlights of the Year

Jointly with Univ. Of Kaiserslautern (C. Eder), we have released a new open source C library for linear algebra dedicated to Gröbner bases computations (see http://www-polsys.lip6.fr/~jcf/Software/index.html). This new library opens the door to high performance applications

- The library is specialized in reducing matrices generated during Gröbner bases computations. Optimizing this reduction step is crucial for the overall computation.
- Our approach takes even more advantage of the very special structure (quasi unit-triangular sparse matrices with patterns in the data)
- We also reduce the number of operations, in a parallel friendly fashion, by changing the order of the operations in the elimination.
- We present experimental results for sequential and parallel computations on NUMA architectures. We also get good scaling up until 32 (non hyper-threaded) cores: we have speed-ups around 14 or 16.

6.2. Fundamental algorithms and structured polynomial systems

6.2.1. Sparse Gröbner Bases

Sparse elimination theory is a framework developped during the last decades to exploit monomial structures in systems of Laurent polynomials. Roughly speaking, this amounts to computing in a *semigroup algebra*, *i.e.* an algebra generated by a subset of Laurent monomials. In order to solve symbolically sparse systems, we introduce *sparse Gröbner bases*, an analog of classical Gröbner bases for semigroup algebras, and we propose sparse variants of the F_5 and FGLM algorithms to compute them.

In the case where the generating subset of monomials corresponds to the points with integer coordinates in a normal lattice polytope $\mathcal{P} \subset \mathbb{R}^n$ and under regularity assumptions, we prove in [19] complexity bounds which depend on the combinatorial properties of $\mathcal{P}$. These bounds yield new estimates on the complexity of solving 0-dim systems where all polynomials share the same Newton polytope (*unmixed case*). For instance, we generalize the bound $\min(n_1, n_2) + 1$ on the maximal degree in a Gröbner basis of a 0-dim. Bilinear system with blocks of variables of sizes (n_1, n_2) to the multihomogeneous case: $n + 2 - \max_i(\lceil (n_i+1)/d_i \rceil)$. We also propose a variant of Fröberg's conjecture which allows us to estimate the complexity of solving overdetermined sparse systems.

Moreover, our prototype "proof-of-concept" implementation shows large speed-ups (more than 100 for some examples) compared to optimized (classical) Gröbner bases software.

6.2.2. Gröbner bases for weighted homogeneous systems

Solving polynomial systems arising from applications is frequently made easier by the structure of the systems. Weighted homogeneity (or quasi-homogeneity) is one example of such a structure: given a system of weights $W = (w_1, \cdots, w_n)$, W-homogeneous polynomials are polynomials which are homogeneous w.r.t the weighted degree $\deg_W(X_1^{\alpha_1}, \cdots, X_n^{\alpha_n}) = \sum w_i \alpha_i$.

Gröbner bases for weighted homogeneous systems can be computed by adapting existing algorithms for homogeneous systems to the weighted homogeneous case. In [29], we show that in this case, the complexity estimate for Algorithm F5 $\left(\binom{n+d_{\max}-1}{d_{\max}}^\omega\right)$ can be divided by a factor $(\prod w_i)^\omega$. For zero-dimensional systems, the complexity of Algorithm FGLMnD^ω (where D is the number of solutions of the system) can be divided by the same factor $(\prod w_i)^\omega$. Under genericity assumptions, for zero-dimensional weighted homogeneous systems of W-degree $(d_1, \cdots, d_n)$, these complexity estimates are polynomial in the weighted Bézout bound $\prod_{i=1}^n d_i / \prod_{i=1}^n w_i$.

Furthermore, the maximum degree reached in a run of Algorithm F5 is bounded by the weighted Macaulay bound $\sum (d_i - w_i) + w_n$, and this bound is sharp if we can order the weights so that $w_n = 1$. For overdetermined semi-regular systems, estimates from the homogeneous case can be adapted to the weighted case.

We provide some experimental results based on systems arising from a cryptography problem and from polynomial inversion problems. They show that taking advantage of the weighted homogeneous structure yields substantial speed-ups, and allows us to solve systems which were otherwise out of reach.

6.2.3. Computing necessary integrability conditions for planar parametrized homogeneous potentials

Let $V \in \mathbb{Q}(i)(\mathbf{a}_1, \cdots, \mathbf{a}_n)(\mathbf{q}_1, \mathbf{q}_2)$ be a rationally parametrized planar homogeneous potential of homogeneity degree $k \neq -2, 0, 2$. In [12], we design an algorithm that computes polynomial *necessary* conditions on the parameters $(\mathbf{a}_1, \cdots, \mathbf{a}_n)$ such that the dynamical system associated to the potential V is integrable. These conditions originate from those of the Morales-Ramis-Simó integrability criterion near all Darboux points and make use of Gröbner bases algorithms. The implementation of the algorithm allows to treat applications that were out of reach before, for instance concerning the non-integrability of polynomial potentials up to degree 9. Another striking application is the first complete proof of the non-integrability of the *collinear three body problem*.

6.3. Solving Polynomial Systems over the Reals and Applications

6.3.1. Exact algorithms for polynomial optimization

Let $f, f_1, ..., f_s$ be n-variate polynomials with rational coefficients of maximum degree D and let V be the set of common complex solutions of $\mathbf{F} = (f_1, ..., f_s)$. In [7], we give an algorithm which, up to some regularity assumptions on $\mathbf{F}$, computes an *exact* representation of the global infimum $f^{\bigstar}$ of the restriction of the map $x \to f(x)$ to $V \cap \mathbb{R}^n$, i.e. a univariate polynomial vanishing at $f^{\bigstar}$ and an isolating interval for $f^{\bigstar}$. Furthermore, it decides whether $f^{\bigstar}$ is reached and if so, it returns $x^{\bigstar} \in V \cap \mathbb{R}^n$ such that $f(x^{\bigstar}) = f^{\bigstar}$.

This algorithm is *probabilistic*. It makes use of the notion of polar varieties. Its complexity is essentially *cubic* in $(sD)^n$ and linear in the complexity of evaluating the input. This fits within the best known *deterministic* complexity class $D^{O(n)}$.

We report on some practical experiments of a first implementation that is available as a MAPLE package. It appears that it can tackle global optimization problems that were unreachable by previous exact algorithms and can manage instances that are hard to solve with purely numeric techniques. As far as we know, even under the extra genericity assumptions on the input, it is the first probabilistic algorithm that combines practical efficiency with good control of complexity for this problem.

It is known that point searching in basic semialgebraic sets and the search for globally minimal points in polynomial optimization tasks can be carried out using $(s\,d)^{O(n)}$ arithmetic operations, where n and s are the numbers of variables and constraints and d is the maximal degree of the polynomials involved.
Subject to certain conditions, we associate in [2] to each of these problems an intrinsic system degree which becomes in worst case of order $(n\,d)^{O(n)}$ and which measures the intrinsic complexity of the task under consideration.
We design non-uniform deterministic or uniform probabilistic algorithms of intrinsic, quasi-polynomial complexity which solve these problems.

6.3.2. Algorithms for answering connectivity queries

Let $\mathbf{R}$ be a real closed field and $\mathbf{D} \subset \mathbf{R}$ an ordered domain. In [4], we give an algorithm that takes as input a polynomial $Q \in \mathbf{D}[X_1, ..., X_k]$, and computes a description of a roadmap of the set of zeros, $\mathrm{Zer}(Q, \mathbf{R}^k)$, of Q in $\mathbf{R}^k$. The complexity of the algorithm, measured by the number of arithmetic operations in the ordered domain $\mathbf{D}$, is bounded by $D^{O(k\sqrt{k})}$, where $D = \deg(Q) \geq 2$. As a consequence, there exist algorithms for computing the number of semi-algebraically connected components of a real algebraic set, $Z(Q, \mathbf{R}^n)$, whose complexity is also bounded by $D^{O(n\sqrt{n})}$, where $D = \deg(Q) \geq 2$. The best previously known algorithm for constructing a roadmap of a real algebraic subset of $\mathbf{R}^n$ defined by a polynomial of degree D has complexity $D^{O(n^2)}$.

In [36], we provide a probabilistic algorithm which computes roadmaps for smooth and bounded real algebraic sets such that the output size and the running time are polynomial in $(nD)^{n\log(n)}$. More precisely, the running time of the algorithm is essentially subquadratic in the output size. Even under these extra assumptions, it is the first roadmap algorithm with output size and running time polynomial in $(nD)^{n\log(n)}$.

6.3.3. Nearly Optimal Refinement of Real Roots of a Univariate Polynomial

In [33], we consider the following problem. 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 $O(d^2\tau + dL)$. This substantially decreases the known bound $O(d^3 + d^2L)$. Furthermore we readily extend our algorithm to support the same complexity bound for the refinement of r real roots, for any $r \leq d$, by incorporating the known efficient algorithms for multipoint polynomial evaluation. The main ingredient for the latter ones is an efficient algorithm for (approximate) polynomial division; we present a variation based on structured matrices computation with quasi-optimal Boolean complexity.

6.3.4. Accelerated Approximation of the Complex Roots of a Univariate Polynomial

Highly efficient and even nearly optimal algorithms have been developed for the classical problem of univariate polynomial root-finding, but this is still an area of active research. By combining some powerful techniques developed in this area we devise in [20] new nearly optimal algorithms, whose substantial merit is their simplicity, important for the implementation.

6.3.5. Nearly Optimal Computations with Structured Matrices

In [21], we estimate the Boolean complexity of multiplication of structured matrices by a vector and the solution of nonsingular linear systems of equations with these matrices. We study four basic most popular classes, that is, Toeplitz, Hankel, Cauchy and Vandermonde matrices, for which the cited computational problems are equivalent to the task of polynomial multiplication and division and polynomial and rational multipoint evaluation and interpolation. The Boolean cost estimates for the latter problems have been obtained by Kirrinnis, except for rational interpolation, which we provide now. All known Boolean cost estimates for these problems rely on using Kronecker product. This implies the d-fold precision increase for the d-th degree output, but we avoid such an increase by relying on distinct techniques based on employing FFT. Furthermore we simplify the analysis and make it more transparent by combining the representation of our tasks and algorithms in terms of both structured matrices and polynomials and rational functions. This also enables further extensions of our estimates to cover Trummer's important problem and computations with the popular classes of structured matrices that generalize the four cited basic matrix classes.

6.3.6. Bounds for the Condition Number for Polynomials with Integer Coefficients

In [31], we consider the problem of bounding the condition number of the roots of univariate polynomials and polynomial systems, when the input polynomials have integer coefficients. We also introduce an aggregate version of the condition numbers and we prove bounds of the same order of magnitude as in the case of the condition number of a single root.

6.4. Solving Systems in Finite Fields, Applications in Cryptology and Algebraic Number Theory

6.4.1. Polynomial-Time Algorithms for Quadratic Isomorphism of Polynomials: The Regular Case

Let $\mathbf{f} = (f_1, ..., f_m)$ and $\mathbf{g} = (g_1, ..., g_m)$ be two sets of $m \geq 1$ nonlinear polynomials in $\mathbb{K}[x_1, ..., x_n]$ ($\mathbb{K}$ being a field). In [25], we consider the computational problem of finding – if any – an invertible transformation on the variables mapping $\mathbf{f}$ to $\mathbf{g}$. The corresponding equivalence problem is known as *Isomorphism of Polynomials with one Secret* (IP1S) and is a fundamental problem in multivariate cryptography. Amongst its applications, we can cite Graph Isomorphism (GI) which reduces to equivalence of cubic polynomials with respect to an invertible linear change of variables, according to Agrawal and Saxena. The main result of our work is a randomized polynomial-time algorithm for solving IP1S for quadratic instances, a particular case of importance in cryptography.

To this end, we show that IP1S for quadratic polynomials can be reduced to a variant of the classical module isomorphism problem in representation theory. We show that we can essentially *linearize* the problem by reducing quadratic-IP1S to test the orthogonal simultaneous similarity of symmetric matrices; this latter problem was shown by Chistov, Ivanyos and Karpinski (ISSAC 1997) to be equivalent to finding an invertible matrix in the linear space $\mathbb{K}^{n\times n}$ of $n \times n$ matrices over $\mathbb{K}$ and to compute the square root in a certain representation in a matrix algebra. While computing square roots of matrices can be done efficiently using numerical methods, it seems difficult to control the bit complexity of such methods. However, we present exact and polynomial-time algorithms for computing a representation of the square root of a matrix in $\mathbb{K}^{n\times n}$, for various fields (including finite fields), as a product of two matrices. Each coefficient of these matrices lie in an extension field of $\mathbb{K}$ of polynomial degree. We then consider #IP1S, the counting version of IP1S for quadratic instances. In particular, we provide a (complete) characterization of the automorphism group of homogeneous quadratic polynomials. Finally, we also consider the more general *Isomorphism of Polynomials* (IP) problem where we allow an invertible linear transformation on the variables *and* on the set of polynomials. A randomized polynomial-time algorithm for solving IP when $\mathbf{f} = (x_1^d, ..., x_n^d)$ is presented. From an algorithmic point of view, the problem boils down to factoring the determinant of a linear matrix (*i.e.* a matrix whose components are linear polynomials). This extends to IP a result of Kayal obtained for PolyProj.

6.4.2. A Polynomial-Time Key-Recovery Attack on MQQ Cryptosystems

In [15], we investigate the security of the family of MQQ public key cryptosystems using multivariate quadratic quasigroups (MQQ). These cryptosystems show especially good performance properties. In particular, the MQQ-SIG signature scheme is the fastest scheme in the ECRYPT benchmarking of cryptographic systems (eBACS). We show that both the signature scheme MQQ-SIG and the encryption scheme MQQ-ENC, although using different types of MQQs, share a common algebraic structure that introduces a weakness in both schemes. We use this weakness to mount a successful polynomial time key-recovery attack. Our key-recovery attack finds an equivalent key using the idea of so-called good keys that reveals the structure gradually. In the process we need to solve a MinRank problem that, because of the structure, can be solved in polynomial-time assuming some mild algebraic assumptions. We highlight that our theoretical results work in characteristic 2 which is known to be the most difficult case to address in theory for MinRank attacks. Also, we emphasize that our attack works without any restriction on the number of polynomials removed from the public-key, that is, using the minus modifier. This was not the case for previous MinRank like-attacks against MQ schemes. From a practical point of view, we are able to break an MQQ-SIG instance of 80 bits security in less than 2 days, and one of the more conservative MQQ-ENC instances of 128 bits security in little bit over 9 days. Altogether, our attack shows that it is very hard to design a secure public key scheme based on an easily invertible MQQ structure.

6.4.3. Algebraic Cryptanalysis of a Quantum Money Scheme – The Noise-Free Case

In [13], we investigate the Hidden Subspace Problem (HSP_q) over $\mathbb{F}_q$:

Input : $p_1, ..., p_m, q_1, ..., q_m \in \mathbb{F}_q[x_1, ..., x_n]$ of degree $d \geq 3$ (and $n \leq m \leq 2n$).
Find : a subspace $A \subset \mathbb{F}_q^n$ of dimension $n/2$ (n is even) such that

$$p_i(A) = 0 \ \forall i \in \{1, ..., m\} \text{ and } q_j(A^\perp) = 0 \ \forall j \in \{1, ..., m\},$$

where $A^\perp$ denotes the orthogonal complement of A with respect to the usual scalar product in $\mathbb{F}_q$.

This problem underlies the security of the first public-key quantum money scheme that is proved to be cryptographically secure under a non quantum but classic hardness assumption. This scheme was proposed by S. Aaronson and P. Christiano at STOC'12. In particular, it depends upon the hardness of HSP_2. More generally, Aaronson and Christiano left as an open problem to study the security of the scheme for a general field $\mathbb{F}_q$. We present a randomized polynomial-time algorithm that solves the HSP_q for $q > 2$ with success probability $\approx 1 - 1/q$. So, the quantum money scheme extended to $\mathbb{F}_q$ is not secure. Finally, based on experimental results and a structural property of the polynomials that we prove, we conjecture that there is also a randomized polynomial-time algorithm solving the HSP_2 with high probability. To support our theoretical results, we also present several experimental results confirming that our algorithms are very efficient in practice. We emphasize that Aaronson and Christiano propose a non-noisy and a noisy version of the public-key quantum money scheme. The noisy version of the quantum money scheme remains secure.

6.4.4. Algebraic Algorithms for LWE Problems

In [23], we analyse the complexity of algebraic algorithms for solving systems of linear equations with *noise*. Such systems arise naturally in the theory of error-correcting codes as well as in computational learning theory. More recently, linear systems with noise have found application in cryptography. The *Learning with Errors* (LWE) problem has proven to be a rich and versatile source of innovative cryptosystems, such as fully homomorphic encryption schemes. Despite the popularity of the LWE problem, the complexity of algorithms for solving it is not very well understood, particularly when variants of the original problem are considered. Here, we focus on and generalise a particular method for solving these systems, due to Arora & Ge, which reduces the problem to non-linear but noise-free system solving. Firstly, we provide a refined complexity analysis for the original Arora-Ge algorithm for LWE. Secondly, we study the complexity of applying algorithms for computing Gröbner basis, a fundamental tool in computational commutative algebra, to solving Arora-Ge-style systems of non-linear equations. We show positive and negative results. On the one hand, we show that the use of Gröbner bases yields an exponential speed-up over the basic Arora-Ge approach. On the other hand, we give a negative answer to the natural question whether the use of such techniques can yield a subexponential algorithm for the LWE problem. Under a mild algebraic assumption, we show that it is highly unlikely that such an improvement exists. We also consider a variant of LWE known as BinaryError-LWE introduced by Micciancio and Peikert recently. By combining Gröbner basis algorithms with the Arora-Ge modelling, we show under a natural algebraic assumption that BinaryError-LWE can be solved in subexponential time as soon as the number of samples is quasi-linear. We also derive precise complexity bounds for BinaryError-LWE with $m = O(n)$, showing that this new approach yields better results than best currently-known generic (exact) CVP solver as soon as $m/n \geq 6.6$. More generally, our results provide a good picture of the hardness degradation of BinaryError-LWE for various number of samples.. This addresses an open question from Micciancio and Peikert. Whilst our results do not contradict the hardness results obtained by Micciancio and Peikert, they should rule out BinaryError-LWE for many cryptographic applications. The results in this work depend crucially on the assumption the algebraic systems considered systems are not easier and not harder to solve than a random system of equations. We have verified experimentally such hypothesis. We also have been able to prove formally the assumptions is several restricted situations. We emphasize that these issues are highly non-trivial since proving our assumptions in full generality would allow to prove a famous conjecture in commutative algebra known as Fröberg's Conjecture.

6.4.5. Practical Cryptanalysis of a Public-Key Encryption Scheme Based on New Multivariate Quadratic Assumptions

In [10], 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 the simplest of our attacks. As a proof of concept, we present 3 practical attacks against all the parameters proposed by Huang, Liu and Yang. With the most efficient attack, we have been able to recover the private-key in roughly 5 minutes for the first challenge (i.e. Case 1) proposed by HLY and less than 30 minutes for the second challenge (i.e. Case 2).

6.4.6. Lazy Modulus Switching for the BKW Algorithm on LWE

Some recent constructions based on LWE do not sample the secret uniformly at random but rather from some distribution which produces small entries. The most prominent of these is the binary-LWE problem where the secret vector is sampled from $\{0,1\}^*$ or $\{-1,0,1\}^*$. In [9], we present a variant of the BKW algorithm for binary-LWE and other small secret variants and show that this variant reduces the complexity for solving binary-LWE. We also give estimates for the cost of solving binary-LWE instances in this setting and demonstrate the advantage of this BKW variant over standard BKW and lattice reduction techniques applied to the SIS problem. Our variant can be seen as a combination of the BKW algorithm with a lazy variant of modulus switching which might be of independent interest.

In [1], 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.

6.4.7. Algebraic Attack against Variants of McEliece with Goppa Polynomial of a Special Form

In [17], we present a new algebraic attack against some special cases of Wild McEliece Incognito, a generalization of the original McEliece cryptosystem. This attack does not threaten the original McEliece cryptosystem. We prove that recovering the secret key for such schemes is equivalent to solving a system of polynomial equations whose solutions have the structure of a usual vector space. Consequently, to recover a basis of this vector space, we can greatly reduce the number of variables in the corresponding algebraic system. From these solutions, we can then deduce the basis of a GRS code. Finally, the last step of the cryptanalysis of those schemes corresponds to attacking a McEliece scheme instantiated with particular GRS codes (with a polynomial relation between the support and the multipliers) which can be done in polynomial-time thanks to a variant of the Sidelnikov-Shestakov attack. For Wild McEliece & Incognito, we also show that solving the corresponding algebraic system is notably easier in the case of a non-prime base field $\mathbb{F}_q$. To support our theoretical results, we have been able to practically break several parameters defined over a non-prime base field $q \in \{9, 16, 25, 27, 32\}, t < 7$, extension degrees $m \in \{2, 3\}$, security level up to 2^{129} against information set decoding in few minutes or hours.

6.4.8. Folding Alternant and Goppa Codes with Non-Trivial Automorphism Groups

The main practical limitation of the McEliece public-key encryption scheme is probably the size of its key. A famous trend to overcome this issue is to focus on subclasses of alternant/Goppa codes with a non trivial automorphism group. Such codes display then *symmetries* allowing compact parity-check or generator matrices. For instance, a key-reduction is obtained by taking *quasi-cyclic* (QC) or *quasi-dyadic* (QD) alternant/Goppa codes. We show in [6], [18], [28] that the use of such *symmetric* alternant/Goppa codes in cryptography introduces a fundamental weakness. It is indeed possible to reduce the key-recovery on the original symmetric public-code to the key-recovery on a (much) smaller code that has not anymore symmetries. This result is obtained thanks to a new operation on codes called *folding* that exploits the knowledge of the automorphism group. This operation consists in adding the coordinates of codewords which belong to the same orbit under the action of the automorphism group. The advantage is twofold: the reduction factor can be as large as the size of the orbits, and it preserves a fundamental property: folding the dual of an alternant (*resp*. Goppa) code provides the dual of an alternant (*resp*. Goppa) code. A key point is to show that all the existing constructions of alternant/Goppa codes with symmetries follow a common principal of taking codes whose support is globally invariant under the action of affine transformations (by building upon prior works of T. Berger and A. Dür). This enables not only to present a unified view but also to generalize the construction of QC, QD and even *quasi-monoidic* (QM) Goppa codes. All in all, our results can be harnessed to boost up any key-recovery attack on McEliece systems based on symmetric alternant or Goppa codes, and in particular algebraic attacks.

6.4.9. Rounding and Chaining LLL: Finding Faster Small Roots of Univariate Polynomial Congruences

In a seminal work at EUROCRYPT '96, Coppersmith showed how to find all small roots of a univariate polynomial congruence in polynomial time: this has found many applications in public-key cryptanalysis and in a few security proofs. However, the running time of the algorithm is a high-degree polynomial, which limits experiments: the bottleneck is an LLL reduction of a high-dimensional matrix with extra-large coefficients. We present in [11] the first significant speedups over Coppersmith's algorithm. The first speedup is based on a special property of the matrices used by Coppersmith's algorithm, which allows us to provably speed up the LLL reduction by rounding, and which can also be used to improve the complexity analysis of Coppersmith's original algorithm. The exact speedup depends on the LLL algorithm used: for instance, the speedup is asymptotically quadratic in the bit-size of the small-root bound if one uses the Nguyen-Stehlé L2 algorithm. The second speedup is heuristic and applies whenever one wants to enlarge the root size of Coppersmith's algorithm by exhaustive search. Instead of performing several LLL reductions independently, we exploit hidden relationships between these matrices so that the LLL reductions can be somewhat chained to decrease the global running time. When both speedups are combined, the new algorithm is in practice hundreds of times faster for typical parameters.

6.4.10. Symmetrized summation polynomials: using small order torsion points to speed up elliptic curve index calculus

Decomposition-based index calculus methods are currently efficient only for elliptic curves E defined over non-prime finite fields of very small extension degree n. This corresponds to the fact that the Semaev summation polynomials, which encode the relation search (or "sieving"), grows over-exponentially with n. Actually, even their computation is a first stumbling block and the largest Semaev polynomial ever computed is the 6-th. Following ideas from Faugère, Gaudry, Huot and Renault, our goal is to use the existence of small order torsion points on E to define new summation polynomials whose symmetrized expressions are much more compact and easier to compute. This setting allows to consider smaller factor bases, and the high sparsity of the new summation polynomials provides a very efficient decomposition step. In [16], the focus is on 2-torsion points, as it is the most important case in practice. We obtain records of two kinds: we successfully compute up to the 8-th symmetrized summation polynomial and give new timings for the computation of relations with degree 5 extension fields.

6.4.11. Sub-cubic Change of Ordering for Gröner Basis: A Probabilistic Approach

The usual algorithm to solve polynomial systems using Gröbner bases consists of two steps: first computing the DRL Gröbner basis using the F5 algorithm then computing the LEX Gröbner basis using a change of ordering algorithm. When the Bézout bound is reached, the bottleneck of the total solving process is the change of ordering step. For 20 years, thanks to the FGLM algorithm the complexity of change of ordering is known to be cubic in the number of solutions of the system to solve. We show in [14] that, in the generic case or up to a generic linear change of variables, the multiplicative structure of the quotient ring can be computed with no arithmetic operation. Moreover, given this multiplicative structure we propose a change of ordering algorithm for Shape Position ideals whose complexity is polynomial in the number of solutions with exponent ω where $2 \leq \omega < 2.3727$ is the exponent in the complexity of multiplying two dense matrices. As a consequence, we propose a new Las Vegas algorithm for solving polynomial systems with a finite number of solutions by using Gröbner basis for which the change of ordering step has a sub-cubic (i.e. with exponent ω) complexity and whose total complexity is dominated by the complexity of the F5 algorithm. In practice we obtain significant speedups for various polynomial systems by a factor up to 1500 for specific cases and we are now able to tackle some instances that were intractable.

PROSECCO Project-Team

6. New Results

6.1. Highlights of the Year

This year, we published 17 articles in international peer-reviewed journals and conferences, including papers in prestigious conferences such as POPL (2 papers) and all the top conferences in computer security: IEEE S&P Oakland (2 papers), CRYPTO, ACM CCS, NDSS, and Financial Cryptography. Our papers in these top venues (discussed later in New Results) serve as highlights of our research during the year. In addition to these papers, we published 1 PhD thesis and several technical reports.

We released updates to miTLS, ProVerif, CryptoVerif, and started working on a brand-new version of F*. We discovered serious vulnerabilities in a number of TLS libraries, web browsers, and web servers, resulting in 6 published CVEs, and over a dozen software updates based on our recommendations in widely used software such as Firefox, Chrome, Internet Explorer, Safari, OpenSSL, Java, and Mono.

We organized a winter school "The Joint EasyCrypt-F*-CryptoVerif School 2014" which attracted industrial researchers, academics, and students from around the world. Over 75 students learned to use cryptographic verification tools from instructors at Inria, IMDEA, and Microsoft Research. Two of the tools: CryptoVerif and F* are being developed in collaboration with Inria.

If we were to choose one research theme as our highlight of the year, it would be our activities surrounding Transport Layer Security (TLS):

- At CRYPTO 2014, we published a detailed cryptographic proof of the TLS handshake as implemented in miTLS
- At NDSS 2014, we published a study in the use of X.509 certificates in TLS servers on the web
- At IEEE S&P (Oakland), we published a new attack on the TLS protocol called the *triple handshake*, which affected all TLS libraries and mainstream TLS applications such as web browsers.
- To prevent our attack, we proposed patches to major software libraries as part of responsible disclosure. Our research directly led to security updates for all major web browsers and TLS implementations.
- We also proposed a long-term countermeasure for our attack, the TLS Session Hash extension, which we published as an internet draft and presented at the IETF. This draft is on its way to being a published standard and is already implemented in all major TLS libraries.
- We participated in the design of next version (1.3) of the TLS protocol. We hosted an interim TLS working group meeting in Paris. Our proposals such as the session hash construction are now an integral part of the new design, and we continue consulting on the design and implementation of TLS.

6.2. Verification of Security Protocols in the Symbolic Model

Participants: Bruno Blanchet, Miriam Paiola, Robert Künnemann.

Miriam Paiola wrote and defended her PhD thesis on the verification of security protocols with lists [45].

Robert Künnemann published a paper at the IEEE S&P conference on how to extend symbolic cryptographic protocol verifiers to account for global state [60].

Bruno Blanchet published a tutorial on the protocol verifier PROVERIF [66], as a follow-up to his teaching in the FOSAD'13 summer school last year.

The applied pi calculus is a widely used language for modeling security protocols, including as a theoretical basis of PROVERIF. However, the seminal paper that describes this language (Abadi and Fournet, POPL'01) does not come with proofs, and detailed proofs for the results in this paper were never published. This year, Martin Abadi, Bruno Blanchet, and Cedric Fournet wrote detailed proofs for the main theorems of this paper. This work was also an opportunity to fix a few minor details in the results and to tune the calculus to improve it and make it closer to the input calculus of PROVERIF. We plan to submit this work as a journal paper.

6.3. Verification of Security Protocols in the Computational model

Participants: Bruno Blanchet, David Cadé.

We worked on our computationally-sound protocol verifier CRYPTOVERIF in two directions.

First, this verifier includes a specialized compiler that generates secure implementations of protocols from CRYPTOVERIF specifications. We completed a journal version of the proof that this compiler preserves security, which is to appear in the Journal of Computer Security [48]

Second, Bruno Blanchet extended CRYPTOVERIF with support for equational theories: associativity, commutativity, non-commutative and commutative groups, exclusive or. The goal is to be able to verify protocols that rely on the algebraic properties of groups and exclusive or. The extended tool is available at http://cryptoverif.inria.fr.

6.4. Computationally Complete Symbolic Attacker Models

Participants: Gergei Bana, Hubert Comon-Lundh.

A new approach to computational verification is to define a *computationally complete* symbolic attacker, so that a symbolic proof against this attacker can be shown to imply a computational proof of security. Following this line of inquiry, Gergei Bana and Hubert Comon-Lundh recently published work on proving computational reachability properties using symbolic techniques.

Gergei Bana (along with Hubert Comon-Lundh) published a paper on how to extend this work to prove stronger security properties expressed as equivalences [50]. Hence, the proof techniq can now be used also for properties like anonymity, strong secrecy etc. Besides being able to prove such properties, another advantage of this extension is that modern security properties of cryptographic primitives are also formulated in terms of indistinguishability, which makes it easier to translate the security properties cryptographers define to our language than before.

Using the computationally complete symbolic attacker, writing up a full, computationally sound proof (and identifying new attacks) for the NSL protocol when agents can run both roles, including running sessions with themselves. The proof is first attempted without any assumption other than that the encryption is CCA2 and that honest names are assigned at the beginning (that is, absolutely nothing about parsing: triples may be independent from pairs, pairing the projection of pairs may not give back the original item etc.). Along the way, we identified new attacks absent of some necessary parsing properties that implementations may not satisfy in general. Then with these additional parsing properties added to the properties satisfied by the implementation, we verified the protocol, namely secrecy, authentication and agreement. The project included graphical representation of the proof steps and the attacks. Type-flaw attacks that can be found in the literature have been reproduced this way, but a number of other attacks have also be revealed that cannot be found with the Dolev-Yao technique, and have not been found by other computational techniques either, although they are realistic. This is joint work with Pedro Adao of IST Lisbon. We hope to publish parts of this work to illustrate proving strategies. The current state of the writeup is available at http://prosecco.gforge.inria.fr/personal/gebana/nsl-long-both-roles.pdf

6.5. Authentication Attacks against Transport Layer Security

Participants: Karthikeyan Bhargavan [correspondant], Antoine Delignat-Lavaud, Cedric Fournet [Microsoft Research], Markulf Kohlweiss [Microsoft Research], Alfredo Pironti, Pierre-Yves Strub [IMDEA].

We discovered an important client impersonation attack on the Transport Layer Security protocol called the *triple handshake attack.* The attack is on the standard protocol and hence all compliant implementations were potentially at risk. Hence, we systematically followed responsible disclosure by notifying all major web browsers and TLS implementors, and then working with the TLS working group to design a countermeasure. The research results of this work were published at IEEE S&P [53].

To TLS implementors, we proposed short-term countermeasures that mitigated our attack, leading to security updates to all major web browsers: Google Chrome (CVE-2013-6628), Mozilla Firefox (CVE-2014-1491), Internet Explorer (CVE-2014-1771), Apple Safari (CVE-2014-1295), as well as to non-browser TLS libraries such as Oracle JSSE (CVE-2014-6457) and RSA BSAFE (CVE-2014-4630). For more details, see http://secure-resumption.com

To the TLS working group, we proposed a new cryptographic construction called the *session hash* that fundamentally alters the cryptographic core of TLS. This construction has now been adopted as a protocol extension to TLS 1.2 and has been integrated into the upcoming TLS 1.3. We expect an IETF standard for this construction to be published in early 2015.

While the triple handshake attacks primarily affect client-authentication, server authentication in HTTPS (HTTP over TLS) primarily relies on X.509 public key certificates. Antoine Delignat-Lavaud along with co-authors at Microsoft research published a paper at NDSS 2015 on a large-scale study of the Web PKI: how certificates are issued and used on the web [56]. Our work uncovered many unsafe practices and suggested best practices and new security policies.

Antoine Delignat-Lavaud also showed how the unsafe sharing of certificates across multiple websites could be exploited to fully compromise the same origin policty for websites, using an vulnerability called virtual host confusion. These results were discussed in a talk at BlackHat USA: for details see http://bh.ht.vc. A research paper on these attacks is forthcoming at WWW'2015.

6.6. A Verified Reference Implementation of Transport Layer Security

Participants: Benjamin Beurdouche [correspondant], Karthikeyan Bhargavan [correspondant], Antoine Delignat-Lavaud, Cedric Fournet [Microsoft Research], Markulf Kohlweiss [Microsoft Research], Alfredo Pironti, Pierre-Yves Strub [IMDEA], Santiago Zanella-Béguelin [Microsoft Research], Jean Karim Zinzindohoue.

Following on from previous work in the miTLS project, we published new versions of miTLS (http://mitls.org) that implemented various protocol extensions including the new session hash extension.

At CRYPTO 2014 [55], we published the first detailed cryptographic proof of an implementation of the TLS Handshake. The implementation consists of about 5000 lines of code and is equipped with about 2500 lines of security annotations written in F7, and a 3000 line EasyCrypt proof.

Currently, we are extending and improving this verified implementation to cover commonly used TLS extensions as well as TLS 1.3, the new version of TLS that we are actively involved in designing. We recently hosted a meeting of the TLS working group at Inria in Paris and are active members of the core working group.

In parallel, we have been analyzing other implementations of TLS and testing them against our implementation, both to ensure interoperability and to uncover bugs. Our analyses have led to the discovery of serious state machine vulnerabilities in many TLS implementations including Oracle JSSE, NSS, OpenSSL, SecureTransport, CyaSSL, Mono, and RSA BSAFE. On our recommendations, all these TLS libraries have issued important security updates in 2014.

6.7. Verified implementations of cryptographic primitives

Participants: Evmorfia-Iro Bartzia, Jean Karim Zinzindohoue, Pierre-Yves Strub, Karthikeyan Bhargavan.

Cryptographic libraries underpin the security of all security protocol implementations. A bug in the implementation of one primitive could enable an attacker to break the security of the full protocol. Hence, establishing the formal correctness of an efficient cryptographic mechanism is a much-desired but still open goal. We are investigating two directions of research towards this goal, specifically in the context of elliptic curve libraries.

Evmorfia-Iro Bartzia and Pierre-Yves Strub are building a Coq library that enables the precise proof of elliptic curve algorithms, and the automatic extraction of verified OCaml code that implements these algorithms. Their most recent result is the formal proof of a non-trivial theorem by Picard: the existence of an isomorphism between an elliptic curve and its Picard group of divisors. This work led to the publication "A formal library for Elliptic Curves in the Coq proof Assistant" and was presented at the ITP 2014 conference [51]. We have also been working on a formal proof of correctness of the GLV algorithm for scalar multiplication in Coq, using the above development and the CoqEal methodology. At present, we have an implementation of the algorithm in the OCaml language and a formal development regarding multiexponentiation, endomorphisms, scalar decomposition and coordinates in both affine and projective spaces. This work is still in progress.

Jean Karim Zinzindohoue and Karthikeyan Bhargavan are investigating the direct verification of implementations of the Curve25519 elliptic curve that is emerging as the preferred new curve for a variety of cryptographic standards, including TLS and the W3C web cryptography API. We use standard program verification tools such as the Frama-C/Why3 verification toolkit for a C implementation of Curve25519 and the F* typechecker for an OCaml implementation of the curve. This work is still in progress.

6.8. Dynamic Security Verification and Testing

Participants: Catalin Hritcu, Arthur Azevedo de Amorim, Zoi Paraskevopoulou, Nikolaos Giannarakis.

We investigated two directions in the runtime security verification of software and hardware systems.

Catalin Hritcu, Arthur Azevedo de Amorim, Nick Giannarakis, and their collaborators at University of Pennsylvania and Portland State University published work on *micro-policies* a generic framework for defining tag-based reference monitors on a simple tagged RISC processor. The framework was formalized and verified in the Coq proof assistant and was used to define and verify micro-policies for dynamic sealing, control-flow integrity, compartmentalization, and memory safety. This work resulted in publications at POPL 2014 [63], ASPLOS 2015 [58], and another paper is in submission.

Catalin Hritcu along with his co-authors worked on a testing framework for security and functional correctness. We published a journal paper about testing noninterference [68] and submitted an ANR JCJC grant pre-proposal on the whole project. Catalin Hritcu also worked with an intern Zoe Paraskevopoulou on this topic, who successfully defended her thesis at NTU Athens. We plan to publish a polished version of that in the near future.

6.9. Verified Security for Web Applications

Participants: Karthikeyan Bhargavan [correspondant], Chetan Bansal [Microsoft], Antoine Delignat-Lavaud, Sergio Maffeis [Imperial College London].

Karthikeyan Bhargavan, Antoine Delignat-Lavaud, and co-authors published a tutorial on Defensive JavaScript, a typed subset of JavaScript that is designed to be used for security-critical components such as cryptographic libraries that may be deployed within untrusted web pages. This tutorial was published as a follow-up of Karthikeyan Bhargavan's lectures at the FOSAD'13 summer school [65].

Karthikeyan Bhargavan, Antoine Delignat-Lavaud, and co-authors also published a journal version of their work on the WebSpi web security modeling library [47], one of the few formal models that captures the detailed security assumptions of various web mechanisms.

Karthikeyan Bhargavan along with collaborators at Microsoft Research published a paper at POPL 2014 on TS*: a new gradual type system for a large subset of JavaScript [47]. We showed how to compile and safely deploy well-typed TS* programs as standard JavaScript in websites. Such programs preserve their types even if other code running on the website is malicious. Our work was used as a basis for further work on the TypeScript compiler and typechecker developed at Microsoft.

6.10. Electronic Voting and Auctions

Participants: Benjamin Smyth [correspondant], Elizabeth Quaglia, Adam Mccarthy, David Bernhard.

Benjamin Smyth continued his work on proving privacy properties of electronic voting protocols. Smyth and Bernhard worked on a new formal definition of ballot secrecy that works even if the bulletin board (used to publish votes) is malicious [69].

Benjamin Smyth, Elizabeth Quaglia, and Adam McCarthy observed that existing electronic voting schemes could be used as core building blocks for electronic auction protocols. Using this link, they build two new e-auction protocols Hawk and Aucitas by building on top of the e-voting protocols Helios and Civitas resp. They prove that their protocols enjoy many desired security properties. This result was published at Financial Cryptography 2014 [61].

SECRET Project-Team

6. New Results

6.1. Highlights of the Year

- Rafael Misoczki's PhD thesis on code-based cryptography (defended in November 2013) has been awarded by the Brazilian Society of Computer Science as the best thesis in computer security.
- *Security analysis of some primitives for authentication and authenticated encryption:* authentication is a major functionality in the vast majority of applications. It is usually implemented by a MAC (message authentication code). The main constructions for MAC are based on hash functions, and include the wide-spread HMAC construction. Gaëtan Leurent, together with Itai Dinur, has presented a new generic attack against HMAC when the underlying hash function follows the Haifa construction. This result points out that the hash function in HMAC has to be chosen very carefully and that some of the main families of hash functions may introduce unexpected weaknesses in the associated MAC. Also, the project-team is involved in a national cryptanalytic effort funded by the ANR which aims at evaluating the security of the recently proposed authenticated encryption schemes.
- *Parallel Repetition of Entangled Games:* In a two-player free game G, two cooperating but non communicating players receive inputs taken from two independent probability distributions. Each of them produces an output and they win the game if they satisfy some predicate on their inputs/outputs. The classical (resp. entangled) value of G is the maximum winning probability when the players are allowed to share classical random bits (resp. a quantum state) prior to receiving their inputs. The n-fold parallel repetition of G consists of n instances of G where the parties receive all the inputs at the same time, produce all the outputs at the same time and must win every instance of G. This work by André Chailloux in collaboration with Giannicola Scarpa establishes that the entangled value of the parallel repetition of G decreases exponentially with n, thereby generalizing to the quantum setting Raz's celebrated parallel repetition theorem which is concerned with the classical value of the game. The main tool for proving this result is the introduction of a new information-theoretic quantity: the superposed information cost.

6.2. 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 in 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 authenticated encryption schemes. Most teams in the research community are actually working on the design and on the analysis (cryptanalysis and optimization of the performance) of such primitives.

6.2.1. 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 analyzed. Most of our work in this area is related to an ANR Project named BLOC. Our recent results then mainly concern either the analysis and design of lightweight block ciphers, or the in-depth study of the security of the block cipher standard AES.

Recent results:

- Cryptanalysis of several recently proposed lightweight block ciphers. This includes an attack against the full cipher KLEIN-64 [60], an attack against 8 rounds (out of 12) of PRINCE [48], [77], and an attack against Zorro and its variants [74].
- Formalization and generic improvements of impossible differential cryptanalysis: this type of attacks, even if extensively used, remains not fully understood, and it appears that there are numerous applications where mistakes have been discovered or where the attacks lack optimality. Our work then provides a general framework for impossible differential cryptanalysis including a generic complexity analysis of the optimal attack. Using these advances, we have also presented the best known impossible differential attacks against several ciphers including CLEFIA-128, Camellia, LBlock and Simon [46], [76], [75].
- Design of a new family of block ciphers achieving very good software performance, especially on 8-bit microcontrollers. A nice feature of these ciphers is that they offer an optimal resistance against side-channel attacks in the sense that the cost of Boolean masking is minimized [58].
- Design and study of a new construction for low-latency block ciphers, named *reflection ciphers*, which generalizes the so-called α-reflection property exploited in PRINCE. This construction aims at reducing the implementation overhead of decryption on top of encryption [24]
- 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 [18].
- Analysis of the differential and linear properties of the AES Superbox [65].

6.2.2. Authenticated encryption

A limitation of all classical block ciphers is that they aim at protecting confidentiality only, while most applications need both encryption and authentication. These two functionalities are provided by using a block cipher like the AES together with an appropriate mode of operation. However, it appears that the most widely-used mode of operation for authenticated encryption, AES-GCM, is not very efficient for high-speed networks. Also, the security of the GCM mode completely collapses when an IV is reused. These severe drawbacks have then motivated an international competition named CAESAR, partly supported by the NIST, which has been recently launched in order to define some new authenticated encryption schemes [0]. Our work related to this competition is then two-fold: G. Leurent has participated to the design of a CAESAR candidate named SCREAM. Also, the project-team is involved in a national cryptanalytic effort led by the BRUTUS project funded by the ANR which aims at evaluating the security of all CAESAR candidates.

Recent results:

- Submission of a proposal to the CAESAR competition [88], [67].
- Cryptanalysis of three CAESAR candidates: Wheesht [64], π-cipher [90], LAC [69].

6.2.3. Hash functions and MACS

The international research effort related to the selection of the new hash function standard SHA-3 has led to many important results and to a better understanding of the security offered by hash functions. However, hash functions are used in a huge number of applications with different security requirements, and also form the building-blocks of some other primitives, like MACs. In this context, we have investigated the security of some of these constructions, in order to determine whether some particular constructions for hash functions may affect the security of the associated MACs.

Recent results:

- Improved generic attacks against hash-based MAC, including HMAC, when the hash function follows the Haifa construction [55], [33];
- Attack against Streebog, the new Russian hash function standard: we show that the specific instantiation of the Haifa construction used in Streebog makes it weak against second-preimage attacks [59].

6.2.4. Cryptographic properties and construction of appropriate building blocks

The construction of building blocks which guarantee a high resistance against the known attacks is a major topic within our project-team, for stream ciphers, block ciphers and hash functions. The use of such optimal objects actually leads to some mathematical structures which may be at the origin of new attacks. This work involves fundamental aspects related to discrete mathematics, cryptanalysis and implementation aspects. Actually, characterizing the structures of the building blocks which are optimal regarding to some attacks is very important for finding appropriate constructions and also for determining whether the underlying structure induces some weaknesses or not.

For these reasons, we have investigated several families of filtering functions and of S-boxes which are well-suited for their cryptographic properties or for their implementation characteristics. 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.

[0] http://competitions.cr.yp.to/caesar.html

Recent results:

- Study of the algebraic properties (e.g. the algebraic degree) of the inverses of APN power permutations [19].
- Study of the cryptographic properties, including the degree, the differential uniformity and the size of the image set of permutations of the form $x \mapsto x^s + \gamma \mathsf{Tr}(x^t)$ over a finite field of characteristic two [15]. Since these functions are obtained by slightly modifying a power function, they share similar interesting implementation properties but do not present the weaknesses coming from their structure. In particular, an infinite family of permutations of this form with differential uniformity 4 has been exhibited.
- Definition of an 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 and linear 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 [65].

6.2.5. Symmetric primitives based on lattices

Lattice-based cryptography is an alternative to number-theoretic constructions for public-key cryptography. Lattice-based constructions enjoy a worst-case security reduction to hard lattice problems, and the area is very active, with many new designs offering attractive features.

Recently, this approach has also been used to build symmetric cryptosystems based on lattice problems. While those systems are less efficient than traditional symmetric systems, they are still reasonably efficient, and their security can be related to hard computational problems rather than being only heuristic. In addition, the underlying mathematical structure can offer extra properties such as parallelizability or easy protection against side-channel attacks.

Recent results:

- Design of a family of pseudo-random functions named SPRING which aims to combine the guarantees of security reductions with good performance [44]; implementation of SPRING on FPGA and protection of this hardware implementation against side-channel attacks [47].
- Implementation and side-channel evaluation of the Lapin authentication protocol, based on the LPN problem [57].

6.3. Code-based cryptography

Participants: Julia Chaulet, Adrien Hauteville, Grégory Landais, Nicolas Sendrier, Jean-Pierre Tillich.

Most popular public-key cryptographic schemes rely either on the factorization 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:

- 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 [16], [17].
- Structural cryptanalysis of some variants of McEliece scheme based on alternant codes which have a quasi-cyclic or quasi-dyadic generator matrix [86].
- Cryptanalysis of a variant of the McEliece cryptosystem based on Reed-Solomon codes [16].
- Design of a new variant of McEliece using quasi-cyclic Moderate Density Parity Check (MDPC) codes [39].

6.4. 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 [0], 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 Ministry of Defense.

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 [10].
- Reconstruction of a convolutional code. This reconstruction technique is based on a new method for detecting whether a given binary sequence is a noisy convolutional codeword obtained from an unknown convolutional code [45].
- Reconstruction of the interleaver of a turbo-code from the knowledge of several noisy codewords [63].

6.5. Quantum information theory

Participants: André Chailloux, Anthony Leverrier, Denise Maurice, Jean-Pierre Tillich.

[0] Kerckhoffs stated that principle in a paper entitled *La Cryptographie militaire*, published in 1883.

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 generalize the best classical codes available today.

Our research also covers quantum cryptography where we study the security of efficient protocols for key distribution or coin flipping, in collaboration with experimental groups. More generally, we investigate how quantum theory severely constrains the action of honest and malicious parties in cryptographic scenarios.

Finally, a promising approach to better understand the possibilities of quantum information consists in studying quantum correlations via the notion of nonlocal games, where different parties need to coordinate to answer some questions, but without communicating. The goal here is to analyze the optimal strategies and to quantify the quantum advantage, i.e. how much sharing an entangled quantum state helps compared to sharing classical randomness.

6.5.1. 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.

Recent results:

- 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 [23].
- Design of a decoding algorithm for the family of quantum codes due to Calderbank, Shor and Steane [84].
- Study of quantum error correcting codes with an iterative decoding algorithm [12].
- Error analysis for Boson Sampling, a simplified model for quantum computation [91].

6.5.2. Quantum cryptography

A recent approach to cryptography takes into account that all interactions occur in a physical world described by the laws of quantum physics. These laws put severe constraints on what an adversary can achieve, and allow for instance to design provably secure key distribution protocols. We study such protocols as well as more general cryptographic primitives such as coin flipping with security properties based on quantum theory.

Recent results:

- Composable security proof for a continuous-variable quantum key distribution protocol with coherent states [92], [71], [70].
- Proof of existence of quantum weak coin flipping with arbitrarily small bias [80].
- Experimental implementation of quantum coin flipping [20].
- Study of connections between quantum encodings, non-locality and quantum cryptography [22].

6.5.3. Quantum correlations and nonlocality

Since the seminal work from Bell in the 60's, it has been known that classical correlations obtained via shared randomness cannot reproduce all the correlations obtained by measuring entangled quantum systems. This impossibility is for instance witnessed by the violation of a Bell inequality and is known under the name of "Quantum Nonlocality". In addition to its numerous applications for quantum cryptography, the study of quantum nonlocality and quantum games has become a central topic in quantum information theory, with the hope of bringing new insights to our understanding of quantum theory.

Recent results:

- Proof of parallel repetition of entangled games with exponential decay [52],[82],[32].
- Development of a general framework for the study of quantum correlations with combinatorial tools [35].
- New bounds on the quantum value of nonlocal games with graph-theoretical arguments [51].
- Optimal bounds for parity-oblivious random access codes [50].
- Study of Local Orthogonality, a physical principle upper bounding quantum correlations [21].
- Considerations on the notion of dimension of physical systems and its implications for information processing [14].

TEMPO Team

6. New Results

6.1. Highlights of the Year

The project was created.

6.2. Approximately Timed Simulation

Participants: Vania Joloboff, Shenpeng Wang.

Existing fast simulators such as SimSoC are Loosely Timed. They evaluate the time taken by instructions executed based on an average model. Typically, the clock value is increased by a constant K every N instructions. This is sufficient to test application software with time-outs or to synchronize multicore applications, but it cannot provide a reasonable performance estimate of the embedded software.

To obtain precise parformance estimate, a common practice is to run the software on Cycle Accurate simulators, which provides a performance measure absolutely correct, but take a very long time. This is becoming a bottleneck. In fact, in many cases the software developers need some performance estimate, but do not require cycle precision. 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, but at a simulation speed that is an order of magnitude faster than a cycle accurate one.

It is possible to maintain fast simulation, (though slower than Loosely Timed) whereas predicting reasonably accurate performance. The challenge is to come up with an abstract model of the processor that does not simulate the processor at cycle level but simulate enough to measure elapsed time with good precision. The approach is the following: a modern processor in nominal mode executes at least one instruction per clock cycle. If it does not do so, it is because there is a delay, whether a cache miss, a pipe line stall, etc. If one can simulate enough of the system so that the cause of the delays can be reproduced in the simulation and the delays evaluated, although the details of the system are not reproduced exactly, then the delays estimate may be accurate enough to provide an acceptable margin error. Moreover some of these computation can be done only once, not for each iteration of a loop.

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 estimate the performance by using static analysis of the application control flow graph combined with a minimum of dynamic computation in order to maintain a reasonable simulation speed. We have developed such a fast Approximately Timed ISS, that does not fully simulate 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. This work will be published in 2015 in volume 68 of the WIT Transactions on Information and Communication Technologies (ISBN 978-1-78466-054-3) [10].

6.3. Automated generation of simulator

Participants: Vania Joloboff, Shengpeng Liu.

Developing a simulator for a complete processor represents a lot of work when it is manual coding, and it is error-prone. Several efforts have been made to generate partly or entirely simulators. The dominant approach in the past years has been to use a high level description language of the processor and to generate code with the language compiler, such as LISA [18], MIMOLA [15], EXPRESSION-ADL [17]. But still, the architecture is described manually into the high level language. It is interesting to explore new architectures, but has the same issues as manual coding for simulation of commercial off-the-shelf processors such as ARM an Power architectures. Of course this approach only makes sense if the vendor has at least some semi-formal description of the architecture, which is not the case for Intel, but is the case for ARM, PowerPC and SH.

In order to automatically generate simulators from the vendor specification, we have initiated a new approach: generating the simulator from the specification of the hardware vendor as available from their web site as .pdf document. After a relatively successful initiative using ad-hoc tools, we wanted to pursue this work in a more robust and industrial context, using XML to generate an XML model of the instruction set from the vendor specification in .pdf, formalize some XML model transformations and finally generate directly the simulator code in C++. In addition, we wanted the translator to be architecture independent, not making any assumption during the translation process. However, this work is hitting difficult issues due to the fact that the vendor specification is incomplete and we have to do more manual architecture specific complements to the specification that we anticipated, which seriously weakens the project objective.

6.4. Automated Test

Participants: Mingsong Chen, Haifeng Gu, Xinqian Zhang.

Under the increasing complexity together with the time-to-market pressure, functional validation is becoming a major bottleneck of smart applications running on mobile platforms (e.g., Android, iOS).Unlike traditional software, smartphone applications are reactive and GUI (Graphical User Interface) intensive. The execution of smartphone applications heavily relies on the interactions with users. Manual GUI testing is extremely slow and unacceptably expensive in practice. However, the lack of formal models of user behaviors in the design phase hinders the automation of GUI testing (i.e., test case generation and test evaluation). While thorough test efforts are required to ensure the consistency between user behavior specifications and GUI implementations, few of existing testing approaches can automatically utilize the design phase information to test complex smartphone applications. Based on UML activity diagrams, we propose an automated GUI testing framework called ADAutomation, which supports user behavior modeling, GUI test case generation, and post-test analysis and debugging. The experiments using two industrial smartphone virtual prototypes demonstrate that our approach [16] can not only drastically reduce overall testing time, but also showed to improve the quality of designs.

6.5. SAT based bounded model checking

Participants: Mingsong Chen, Haifeng Gu, Xinqian Zhang.

SAT-based Bounded Model Checking (BMC) is promising for automated generation of directed tests. Due to the state space explosion problem, SAT-based BMC is unsuitable to handle complex properties with large SAT instances or large bounds. In this work, we propose a framework to automatically scale down the SAT falsification complexity by utilizing the decision ordering based learning from decomposed sub-properties. Our framework makes three important contributions: i) it proposes learning-oriented decomposition techniques for complex property falsification, ii) it proposes an efficient approach to accelerate the complex property falsification using the learning from decomposed sub-properties, and iii) it combines the advantages of both property decomposition and property clustering to reduce the overall test generation time. The experimental results [11] using both software and hardware benchmarks demonstrate the effectiveness of our framework.

CLASSIC Project-Team

4. New Results

4.1. Corpus linguistics and Markov substitute processes

Thomas Mainguy and Olivier Catoni studied a new statistical model for natural language modeling, called Markov substitute processes. This model is based on a set of conditional independence properties that are more general than the Markov field assumption. It has connections with context free grammars and forms a collection of exponential families having for this reason nice estimation properties.

4.2. Kernel Principal Component Analysis and spectral clustering

Ilaria Giulini and Olivier Catoni continued their study of dimension free bounds for the estimation of the Gram matrix and more generally for the estimation of the expectation of a random symmetric matrix from an i.i.d. sample. This study, using PAC-Bayes bounds, both leads to new robust estimators with applications to Principal Component Analysis in high of even infinite dimension, and new bounds for the usual empirical Gram matrix estimate. Getting dimension free bounds is important to get new results on Kernel PCA. Applications were also studied to density estimation and to spectral clustering.

GAMMA3 Project-Team

4. New Results

4.1. Serendipity and reduced elements

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

We give a method to constructing Serendipity elements for quads and hexes with full symmetry properties and indicate the reading of their shape functions. We show that, since the degree 5, the Serendipity elements are no longer symmetric but we propose a method resulting in a Lagrange element of degree 5 with full symmetry properties after adding an adequate number of additional nodes.

On the other hand, we show how to guarantee the geometric validity of a given curved element (seen as a patch) of a mesh. This is achieved after writing the patch in a Bézier setting (Bernstein polynomials and control points). In addition, we discuss the case of patch derived from a transfinite interpolation and it is proved that only some of them are Serendipity elements indeed, we return to the same elements as above

We also give a method to constructing Lagrange Serendipity (or reduced) simplices with a detailed description of the triangles of degree 3 and 4. We indicate that higher order triangles are not candidate apart if we impose a restricted polynomial space. We show that a tetrahedron of degree 3 is a candidate while high order elements are not candidate even if a restriction in the polynomial space is considered. In addition, we propose a method for the validation of such elements, in a given mesh, where the validation means the positiveness of the jacobian.

A technical report have been published [29].

4.2. Validity of transfinite and Bézier-Serendipity patches

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

We define generalized transfinite patches for quads and hexes with full symmetry properties. We give a way of constructing those patches by considering the Bézier setting using linear combinations of tensor-product patches of various degree. Those patches are exactly the Bézier-Serendipity patches recently introduced

ASsfor reduced quadrilateral patches, we introduce the so called "Bézier-Serendip" patches. After some recalls about standard Bézier patches, we propose a method to constructing those reduced patches. The corresponding Bernstein polynomials are written by means of linear combinations of the standard Bernstein polynomials. We give a full description of the patches of degree 2, 3, 4 and 5. Since degree 5, the location of the control points is no longer symmetric and to remedy this problem, we propose adding a number of control points which results in *extended* Bézier-Serendip patches. Those reduced patches are in the Bézier framework what the Serendipity elements are in the finite element framework.

A technical report and a paper have been published [30], [17].

4.3. Meshing Strategies and the Impact of Finite Element Quality on the Velocity Field in Fractured Media

Participants: Patrick Laug [correspondant], Géraldine Pichot.

For calculating flow in a fracture network, the mixed hybrid finite element (MHFE) method is a method of choice as it yields a symmetric, positive definite linear system. However, a drawback to this method is its sensitivity to bad aspect ratio elements. For poor-quality triangles, elementary matrices are ill-conditioned, and inconsistent velocity vectors are obtained by inverting these local matrices. In this work, different strategies have been proposed for better reconstruction of the velocity field [21].

4.4. Automatic Mesh Generation of Multiface Models on Multicore Processors

Participant: Patrick Laug [correspondant].

This work started in September 2014, as part of a sabbatical at Polytechnique Montréal. In a previous study, a parallel version of an indirect approach for meshing composite surfaces – also called multiface models – was developed. However, this methodology could be inefficient in practice, as the memory management of most existing CAD (computer aided design) systems use static global caches to save information. In our new approach, CAD queries are fully parallelized, using the Pirate library from Polytechnique Montréal. This library provides a set of C++ classes that implement STEP-compliant B-Rep geometric and topological entities, as well as classes to represent meshes and solutions. By modifying the data structures so that memory caches are local to each face of the geometric model, geometric primitives can efficiently be evaluated in parallel, and performance measurements show significant gains.

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

Participants: Thomas Grosges [correspondant], Dominique Barchiesi, Michael François.

L'utilisation des nombres (pseudo)-aléatoires a pris une dimension importante ces dernières décennies. De nombreuses applications dans le domaine des télécommunications, de la cryptographie, des simulations numériques ou encore des jeux de hasard, ont contribué au développement et à l'usage de ces nombres. Les méthodes utilisées pour la génération de tels nombres (pseudo)-aléatoires proviennent de deux types de processus : physique et algorithmique. Ce projet de recherche a donc pour objectif principal le développement de nouveaux procédés de génération de clés de chiffrement, dits "exotiques", basés sur des processus physiques, multi-échelles, multi-domaines assurant un niveau élevé de sécurité. Deux classes de générateurs basés sur des principes de mesures physiques et des processus mathématiques ont été 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.

4.6. 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.

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).

4.7. 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.

Le contrôle et l'adaptation du maillage lors de la résolution de problèmes couplés ou/et non linéaires reste un problème ouvert et fortement dépendant du type de couplage physique entre les EDP à résoudre. Notre objectif est de développer des modèles stables afin de calculer les dilatations induites par l'absorption d'énergie électromagnétique, par des structures matérielles inférieures au micron. Les structures étudiées sont en particulier des nanoparticules métalliques en condition de résonance plasmon. Dans ce cas, un maximum d'énergie absorbée est attendu, accompagné d'un maximum d'élévation de température et de dilatation. Il faut en particulier développer des modèles permettant de simuler le comportement multiphysique de particules de formes quelconques, pour une gamme de fréquences du laser d'éclairage assez étendue afin d'obtenir une étude spectroscopique de la température et de la dilatation. L'objectif intermédiaire est de pouvoir quantifier la dilatation en fonction de la puissance laser incidente. Le calcul doit donc être dimensionné et permettre finalement des applications dans les domaines des capteurs et de l'ingénierie biomédicale. En effet, ces nanoparticules métalliques sont utilisées à la fois pour le traitement des cancers superficiels par nécrose de tumeur sous éclairage adéquat, dans la fenêtres de transparence cellulaire. Déposées sur un substrat de verre, ces nanoparticules permettent de construire des capteurs utilisant la résonance plasmon pour être plus sensibles (voir projet européen *Nanoantenna* et l'activité génération de nombres aléatoires). Cependant, dans les deux cas, il est nécessaire, en environnement complexe de déterminer la température locale, voire la dilatation de ces nanoparticules, pouvant conduire à un désaccord du capteur, la résonance plasmon étant très sensible aux paramètres géométriques et matériels des nanostructures. Dans ce sens, l'étude permet d'aller plus loin que la "simple" interaction électromagnétique avec la matière du projet européen *Nanoantenna*.

Le travail de l'année 2014 a constitué en la poursuite de l'étude des spécificités de ce type de problème multiphysique pour des structures de forme simple et la mise en place de fonctions test, de référence, pour les développements de maillage adaptatifs pour les modèles multiphysiques éléments finis. Nous espérons pouvoir proposer un projet ANR couplant les points de vue microscopiques et macroscopiques dans les deux années qui viennent.

4.8. Visualization and modification of high-order curved meshes

Participants: Alexis Loyer, Adrien Loseille [correspondant].

During the partnership between Inria and Distene, a new visualization software has been designed. It address the typical operations that are required to quickly assess the newly algorithm developed in the team. In particular, interactive modifications of high-order curved mesh and hybrid meshes has been addressed. The software VIZIR is freely available at https://www.rocq.inria.fr/gamma/gamma/vizir/.

4.9. Mesh adaptive ALE numerical simulation

Participants: Frédéric Alauzet [correspondant], Nicolas Barral, Adrien Loseille.

Running highly accurate numerical simulations with moving geometries is still a challenge today due to their prohibitive cost in CPU time. Using anisotropic mesh adaptation is one way to drastically reduce the size of the problem and to reach the desired accuracy. Previously, we have developed an ALE formulation using mesh connectivity change in order to achieve any complex displacement. Then, this method has been coupled with the unsteady anisotropic mesh adaptation using the fixed-point algorithm. The key point of this work is the use of an ALE metric that takes into account the mesh motion in the metric field definition [24], [14].

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 [25].

4.11. Adaptive multigrid strategies

Participants: Frédéric Alauzet [correspondant], Victorien Menier, Adrien Loseille.

Multigrid is a well known technique used to accelerate the convergence of linear system solutions. Using a multigrid strategy to solve non-linear problems improves the robustness and the convergence of each Newton step, the accelerating overall the whole process. In particular, larger time step can be considered. This of main importance when solving turbulent Navier-Stokes equations on complex geometries. First, we developed the classical multigrid method on non-nested meshes. Then, we have pointed out the similarity between the Full MultiGrid (FMG) algorithm and the mesh adaptation algorithm. We have proposed a new Adaptive Full MultiGrid algorithm which improve the overall robustness of the adaptive process and its overall efficiency [25].

4.12. Metric-orthogonal and metric-aligned mesh adaptation

Participants: Frédéric Alauzet, Victorien Menier, Adrien Loseille [correspondant].

A new algorithm to derive adaptive meshes has been introduced through new cavity-based algorithms. It allows to generate anisotropic surface and volume mesh that are aligned along the eigenvector directions. This allows us to improv the quality of the meshes and to deal naturally with boundary layer mesh generation [19], [27].

MATHERIALS Team

5. New Results

5.1. Electronic structure calculations

Participants: Eric Cancès, Virginie Ehrlacher, David Gontier, Claude Le Bris, Gabriel Stoltz.

In electronic structure calculation as in most of our scientific endeavours, we pursue a twofold goal: placing the models on a sound mathematical grounding, and improving the numerical approaches.

E. Cancès and N. Mourad have mathematically analyzed the density functional perturbation theory, both in the non-degenerate case (that is, when the Fermi level is not an eigenvalue of the Kohn-Sham hamiltonian) and in the degenerate case. They have in particular proved that Wigner's 2n+1 rule holds in both cases. 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 pioneering work of von Barth and Hedin setting up the Kohn-Sham density functional theory for magnetic compounds. He has also extended a previous work by Anantharaman and Cancès, and proved the existence of minimizers for the spin-polarized Kohn-Sham model in the presence of a magnetic field within the local spin density approximation.

E. Cancès has pursued his long-term collaboration with Y. Maday (UPMC) on the numerical analysis of electronic structure models. With L. He (ENPC) and R. Chakir (IFSTTAR), they 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. Together with G. Dusson (UMPC), B. Stamm (UMPC), and M. Vohralík (Inria), they have designed a new post processing method for planewave discretizations of nonlinear Schrödinger equations, and used it to compute sharp *a posteriori* error estimators for both the discretization error and the algorithmic error (convergence threshold in the iterations on the nonlinearity).

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 medium accounting for long-range electrostatics. E. Cancès, Y. Maday (Paris 6), and B. Stamm (Paris 6) have recently introduced a very efficient domain decomposition method for the simulation of large molecules in the framework of the so-called COSMO implicit solvation models. In collaboration with F. Lipparini (UPMC, B. Mennucci (Department of Chemistry, University of Pisa) and J.-P. Picquemal (Paris 6), they have implemented this algorithm in widely used computational software products (Gaussian and Tinker). The extension of this method to other implicit solvation models is work in progress.

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: Thomas Hudson, Frédéric Legoll, Tony Lelièvre, Mathias Rousset, Gabriel Stoltz.

The work of the team in this area is concentrated on two new directions: the sampling of reactive trajectories (where rare events dictate the dynamics of the system), and the computation of average properties of nonequilibrium systems (which completes the more traditional field of expertise associated with techniques to compute free energy differences).

5.2.1. Sampling of reactive trajectories

Finding trajectories for which the system undergoes a significant change is a challenging task since the transition events are typically very rare. Several methods have been proposed in the physics and chemistry literature, and members of the team have undertaken their study in the past years.

A prominent example is the parallel replica method where several replicas of the system evolve on different processors, until one of them undergoes a transition. Several extensions and refinements to the original method were proposed by T. Lelièvre:

- together with D. Aristoff and G. Simpson, he proposed in [7] an adaptation of the Parallel Replica method for Markov chains;
- together with A. Binder and G. Simpson, he introduced in [17] a generalized parallel replica dynamics. The idea is to extend the applicability of the original algorithm by computing on the fly the so-called decorrelation time.

Another class of techniques to compute reactive trajectories is based on splitting techniques. C.E. Bréhier, T. Lelièvre and M. Rousset have performed in [21] an analysis of the Adaptive Multilevel Splitting algorithm, which is a rare event simulation method where several replicas are evolved concurrently, and selected to favor exploration in a given direction. The computational cost of the algorithm is studied in details in the limit of a large number of replicas.

5.2.2. Nonequilibrium systems

G. Stoltz, together with G. Pavliotis (Imperial College) and Rémi Joubaud, studied in [27] the response of equilibrium systems evolving according to a Langevin dynamics, to external, space-time dependent forcings. In particular, they found out that, even if the external forcing is periodic in time and space with a vanishing space-time average, the systems in general evolves with a non-zero average velocity. It may even be the case that the average velocity is in the direction opposite to the average forcing (when the latter is non-zero), which can be seen as an example of negative mobility. The behavior of the system over diffusive time scales (in the reference frame obtained by removing the average velocity) is also studied, for arbitrary forcing strengths. This work was initiated when G. Pavliotis was a visiting member of the team MATHERIALS.

A numerical analysis of the error arising in the computation of transport coefficients, with an emphasis on mobility and self-diffusion, was provided by M. Fathi, A.A Homman and G. Stoltz in [25] in the case when Metropolis-Hastings algorithms are used to stabilize straightforward discretizations of overdamped Langevin dynamics.

Together with Herbert Spohn (TU München), G. Stoltz has verified the relevance of mode-coupling predictions for the scaling of space time correlations of invariants for one dimensional systems subjected to a non-reversible deterministic dynamics perturbed by an exchange noise [32]. In particular, it has been confirmed that the equilibrium relaxation of the invariants involves two modes, a traveling sound mode and a standing heat mode (related to the energy current and height autocorrelation functions). Both modes exhibit a superdiffusive scaling, of Lévy type for the heat mode, and of KPZ type for the sound mode.

5.2.3. Free energy computations

The topic of free energy computations is still a significant research area of the team. T. Lelièvre and G. Stoltz, together with G. Fort and B. Jourdain, studied the Self-Healing Umbrella Sampling (SHUS) method in [26]. This method is an adaptive biasing method to compute free energies on the fly by appropriately penalizing already visited regions. The convergence of the method relies on a rewriting as a stochastic approximation method with random steps, and can therefore be seen as a variation of the Wang-Landau method. The efficiency of the SHUS algorithm was assessed for a model two-dimensional system in terms of exit times out of a metastable set.

Concerning practical applications, G. Stoltz, together with A.A. Homman, E. Bourasseau, P. Malfreyt, L. Strafella and A. Ghoufi have worked on the computation of surface tension in droplets [10], using alchemical transformations where the droplet volume is artificially varied.

Finally, T. Lelièvre, together with J. Comer, J.C. Gumbart, J. Hénin, A. Pohorille and C. Chipot, wrote a review article on the adaptive biasing force method [9].

5.2.4. Thermodynamic limit

Another work in progress is related to the understanding of the origin of hysteresis in rubber-made materials. When submitted to cyclic deformations, the strain-stress curve of these materials indeed shows a hysteresis behavior, which seems to be independent of the speed of loading.

Some years ago, members of the team have suggested a model, at a mesoscale, to explain this behavior. This model was written in terms of a system made of a finite number of particles. One of the aim of the post-doc of Thomas Hudson, who joined the team in Sept. 2014, is to make progress on that question, and to understand whether a thermodynamic limit of the model previously proposed can be identified.

5.2.5. Reduced models

We propose in [13] a procedure for replacing a complex, reactive potential of REBO type by a simple harmonic approximation, in regions where the system is close to equilibrium. The parameters of the harmonic approximation are chosen so that the phonon spectrum is exactly reproduced. We are currently testing the ability of the so-obtained hybrid model to predict the fracture of graphene.

5.3. Complex fluids

Participants: Sébastien Boyaval, Claude Le Bris, Tony Lelièvre.

Sébastien Boyaval has pursued his research about the mathematical modelling of complex free-surface flows. On the one hand, the numerical investigation of 3D effects with a VOF approach was carried out for multiphase flows in collaboration with the CFSFlow code developers at EPFL [11]. On the other hand, the reduced modelling of viscoelastic effects within Saint-Venant framework was carried out for asymptotically thin layers above rough bottoms [8].

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 very quickly gives reliable results. In particular, the team is interested in two techniques: Proper Generalized Decomposition (greedy algorithms) and Reduced Basis techniques.

Eric Cancès, Virginie Ehrlacher and Tony Lelièvre have extended a greedy algorithm suggested for the resolution of high-dimensional eigenvalue problems in order to approximate the solution of the many-body Schrödinger electronic problem. The main technical difficulty in the extension of these algorithms lies in the antisymmetry of the wavefunction of the electrons. To deal with this difficulty, an approximation of the wavefunction is computed as a sum of Slater determinants, each Slater determinant function being computed in an iterative way.

Virginie Ehrlacher has obtained preliminary encouraging results on greedy algorithms for parametric eigenvalue problems. The method has been applied to the computation of the first buckling mode of a plate in the presence of a defect, the position of the defect playing the role of a parameter entering the eigenvalue problem defining the first buckling mode of the plate.

A new numerical method for the construction of an efficient reduced-order model for the solution of the Vlasov equation, arising in plasma physics or in the modeling of electron transport in semiconductors, has been tested by Damiano Lombardi (REO Inria team) and Virginie Ehrlacher. This method is based on the use of an analytic Lax Pair for the Vlasov equations and is inspired by previous works done on transport equations by Jean-Frederic Gerbeau, Damiano Lombardi and Elisa Schenone. Encouraging preliminary numerical results have been obtained.

5.5. Homogenization and related topics

Participants: Sébastien Brisard, Ludovic Chamoin, Virginie Ehrlacher, Claude Le Bris, Frédéric Legoll, Simon Lemaire, 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. The theoretical results obtained to date are being collected in a series of manuscripts that will be available shortly.

The team 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.

In [28], F. Legoll and W. Minvielle have proposed a variance reduction procedure, based on the control variate technique, to obtain estimates of the apparent homogenized tensor with a smaller statistical error (at a given computational cost) than standard Monte Carlo approaches. The control variate technique is based on using a surrogate model, somewhat in the spirit of a preconditionner. In [28], the surrogate model that is used is inspired by a weakly stochastic approach previously introduced by A. Anantharaman and C. Le Bris to describe periodic models perturbed by rare defects.

In addition, C. Le Bris, F. Legoll and W. Minvielle have investigated the possibility to use another variance reduction technique based on computing the corrector equation only for selected environments. These environments are chosen based on the fact that their statistics in the finite supercell matches the statistics of the materials in the infinite supercell. This method yields an estimator with a smaller variance that standard estimators. Preliminary encouraging numerical results have been obtained.

As pointed out above, the corrector problem is in practice solved on a large bounded domain, often complemented with periodic boundary conditions. Solving that problem can still be challenging, in particular because producing a conforming mesh of realistic heterogeneous microstructures can be a daunting task. In such situations, numerical methods formulated on cartesian grids may be more interesting. These methods can still be Finite Element Methods, or methods in the spirit of that proposed by Moulinec and Suquet in the mid-nineties. In their approach, the corrector problem (a partial differential equation) is reformulated as an equivalent integral equation. This equation can readily be discretized using a Galerkin approach. This leads to numerical schemes that can be implemented as a matrix-free method. In [18], S. Brisard and F. Legoll have reviewed the different variants that have been proposed in the literature along these ideas, and proposed a mathematical analysis of the numerical schemes. This work extends in various directions previous works by S. Brisard.

In somewhat the same vein, Eric Cancès, Virginie Ehrlacher and Frédéric Legoll (in collaboration with Benjamin Stamm, University Paris 6) have worked on alternative methods to approximate the homogenized coefficients of a random stationary material. These methods are alternative to those proposed e.g. by Bourgeat and Piatniski, and which consist in solving a corrector problem on a bounded domain. The method introduced is based on a new corrector problem. This problem is posed on the entire space. In some cases (including the case of randomly located spherical inclusions), it can be recast as an integral equation posed on the surface of the inclusions. The problem can then be efficiently solved via domain decomposition and using spherical harmonics.

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 actually also considered 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 a functional form of the distribution of the microscopic properties is assumed, but with some unknown parameters to determine. In collaboration with A. Obliger and M. Simon, F. Legoll and W. Minvielle have addressed that problem in [29], showing how to determine the unknown parameters of the microscopic distribution on the basis of macroscopic (e.g. homogenized) quantities.

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. That approach has been analyzed and tested on an elliptic problem set on a domain with a huge number of perforations. 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 interplays with the multiscale character of the equation is an unsolved mathematical question worth considering for numerical purposes. In that spirit, C. Le Bris, F. Legoll and F. Madiot have studied several variants of the Multiscale Finite Element Method (MsFEM), specifically designed to address multiscale advection-diffusion problems in the convection-dominated regime. Generally speaking, the idea of the MsFEM is to perform a Galerkin approximation of the problem using specific basis functions, that are precomputed (in an offline stage) and adapted to the problem considered. Several possibilities for the basis functions have been examined (for instance, they may or may not encode the convection field). The various approaches have been compared in terms of accuracy and computational costs.

Most of the numerical analysis studies of the MsFEM are focused on obtaining *a priori* error bounds. In collaboration with L. Chamoin, who is currently in delegation in our team (from ENS Cachan, since September 2014), we have started to work on *a posteriori* error analysis for MsFEM approaches, with the aim to develop error estimation and adaptation tools. We have extended to the MsFEM case an approach that is classical in the computational mechanics community for single scale problems, and which is based on the so-called Constitutive Relation Error (CRE). Once a numerical solution u_h has been obtained, the approach needs additional computations in order to determine a divergence-free field as close as possible to the exact flux $k\nabla u$. In the context of the MsFEM, it is important to be able to do all the expensive computations in an offline stage, independently of the right-hand side. The standard CRE approach thus needs to be adapted to that context, in order to keep that feature that makes it adapted to a multiscale, multi-query context. The preliminary approach that we have introduced already yields promising results.

Still another question investigated in the group is to find an alternative to standard homogenization techniques when these latter are difficult to use in practice. This is the aim of the post-doc of Simon Lemaire, which began in June 2014, and which takes over previous works of the group on the subject. Consider a linear elliptic equation, say in divergence form, with a highly oscillatory matrix coefficient, and assume that this problem is to be solved for a large number of right-hand sides. If the coefficient oscillations are infinitely rapid, the solution can be accurately approximated by the solution to the homogenized problem, where the homogenized coefficient has been evaluated beforehand by solving the corrector problem. If the oscillations are moderately rapid, one can think instead of MsFEM-type approaches to approximate the solution to the

reference problem. However, in both cases, the complete knowledge of the oscillatory matrix coefficient is required, either to build the average model or to compute the multiscale basis. In many practical cases, this coefficient is often only partially known, or merely completely unavailable, and one only has access to the solution of the equation for some loadings. This observation has lead to think about alternative methods, in the following spirit. Is it possible to approximate the reference solution by the solution to a problem with a *constant* matrix coefficient? How can this 'best' constant matrix approximating the oscillatory problem be constructed in an efficient manner?

A preliminary step, following discussion and interaction with A. Cohen, has been to cast the problem as a convex optimization problem. We have then shown that the 'best' constant matrix defined as the solution of that problem converges to the homogenized matrix in the limit of infinitely rapidly oscillatory coefficients. Furthermore, the optimization problem being convex, it can be efficiently solved using standard algorithms. C. Le Bris, F. Legoll and S. Lemaire are currently working on making the resolution of the optimization problem as efficient as possible.

To conclude this section, we mention a project involving V. Ehrlacher, C. Le Bris and F. Legoll, in collaboration with G. Leugering and M. Stingl (Cluster of Excellence, Erlangen-Nuremberg University). This project 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. Materials under consideration are being thought of as microstructured materials composed of steel and void and whose microstructure patterns are constructed as the macroscopic deformation of a reference periodic microstructure. The optimal material (i.e. the best macroscopic deformation) is the deformation achieving the best mechanical response. For a given deformation, we have first chosen to compute the mechanical response using a homogenized model. We are currently aiming at computing the mechanical response at the microscale, using the highly oscillatory model. Model reduction techniques (such as MsFEM, Reduced Basis methods, ...) are then in order, in order to expedite the resolution of the oscillatory problem, which has to be solved at each loop of the optimization algorithm. Current efforts are targeted towards choosing an appropriate model reduction strategy.

5.6. Miscellaneous

Participants: Sébastien Boyaval, Tony Lelièvre, Sébastien Boyaval.

T. Lelièvre together with F. Casenave and A. Ern propose in [24] an extension of the classical reduced basis method in order to extend its range of applicability to black-box codes.

S. Boyaval started investigating new high-order methods on generalized non-conforming meshes in collaboration with Daniele di Pietro [14].

In [31], M. Rousset considers space homogenous Boltzmann kinetic equations in dimension $d \geq 3$ with Maxwell collisions (and without Grad's cut-off). An explicit Markov coupling of the associated conservative stochastic N-particle system is constructed, yielding a N-uniform α-power law trend to equilibrium.

MATHRISK Project-Team

6. New Results

6.1. Highlights of the Year

B. Jourdain and A. Sulem : Guest editors of the special issue "Systemic Risk" of *Statistics and Risk Modeling*, 2014. [27]

The research project "Stochastic Control of Systemic Risk" has been awarded by the scientific council and Professional Fellows of Institut Europlace de Finance (EIF) and Labex Louis Bachelier (December 2014).

Roxana Dumitrescu, PhD student, received the price for collaborative actions during her PhD studies, delivered by Fondation des Sciences Mathématiques de Paris and CASDEN (November 2014).

Pierre Blanc, PhD student, has got the award of "Rising star of quantitative finance" for his talk on a price impact models with an exogeneous (Hawkes) flow of orders [29]. This prize was given by the Global Derivatives conference (Amsterdam, 12-16 May) to indicate the best work among PhD students.

6.2. Liquidity risk

Aurélien Alfonsi and his PhD student Pierre Blanc are working on the optimal execution problem when there are many large traders who modify the price. They consider an Obizhaeva and Wang model for the price impact, and they assume that the flow of market orders generated by the other traders is given by an exogenous process. They have shown that Price Manipulation Strategies (PMS) exist when the flow of order is a compound Poisson process. On the other hand, modeling this flow by a mutually exciting Hawkes process allows them with a particular parametrization to exclude these PMS. Besides, they are able to calculate explicitly the optimal execution strategy within the model [29]. They are now investigating how this model can fit market data.

6.3. Dependence modeling

With his PHD student J. Reygner, B. Jourdain has studied a mean-field version of rank-based models of equity markets, introduced by Fernholz in the framework of stochastic portfolio theory ([38]). When the number of companies grows to infinity, they obtain an asymptotic description of the market in terms of a stochastic differential equation nonlinear in the sense of McKean. The diffusion and drift coefficients depend on the cumulative distribution function of the current marginal law of the capitalizations. Using results on the longtime behavior of such SDEs derived in [66], they 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.

Another approach to handle the question of stochastic modeling in a multidimensional framework consists in dealing with stochastic differential equations that are defined on matrices in order to model either the instantaneous covariance or the instantaneous correlation between the assets.

The research on the estimation of the parameters of a Wishart process has started this year together with the thesis of Clément Rey. A. Alfonsi, A. Kebaier and C. Rey are studying the Maximum Likelihood Estimator for the Wishart processes and in particular its convergence in the ergodic and the non ergodic case.

Correlation issues are crucial in the modeling of volatility. In his thesis, Ould Aly ([77]) proposes a revised version of Bergomi's model for the variance curve which proves to be very tractable for calibration and for the pricing of variance derivatives (see [23]). He also obtains results on the monotonicity of option prices with respect to the correlation between the stock price and the volatility in the Heston model (see [78]).

In [34], [15], L. Abbas-Turki and D. Lamberton study the monotonicity of option prices with respect to cross-asset correlations in a multidimensional Heston model.

Modeling the dependence is not only useful for the equity market. In credit risk, getting a model that describes the dynamic of the joint distribution of a basket of defaults is still a challenge. The Loss Intensity model proposed by Schönbucher allows to fit perfectly the marginal distributions of the number of defaults in a basket. Then, Stochastic Loss Intensity models extend this model and can also in principle fit the marginal distributions. However, these models appear as a non-linear differential equation with jumps. A Alfonsi, C. Labart and J. Lelong have shown that these models are well-defined by using a particles system ([44]). Besides, this particles system gives a very convenient way to run a Monte-Carlo algorithm and to compute expectations in this model. Interacting particle systems are studied by B. Jourdain and his PhD student Julien Reygner in [39], [21].

Application of optimal transport. A. Alfonsi and B. Jourdain study in [43] 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.

6.4. Systemic risk

The mathematical modeling of default contagion, by which an economic shock causing initial losses and default of a few institutions is amplified due to complex linkages, leading to large scale defaults, can be addressed by various techniques, such as network approaches (see in particular [46]), or mean field interaction models [62], [55]. Little has been done so far on the *control* of such systems and A. Sulem has started to contribute on these issues in the framework of random graph models in collaboration with A. Minca (Cornell University) and H. Amini (EPFL). In [22], [31], they consider a financial network described as a weighted directed graph, in which nodes represent financial institutions and edges the exposures between them. Here, the distress propagation is modeled as an epidemics on this graph. They study the optimal intervention of a lender of last resort who seeks to make equity infusions in a banking system prone to insolvency and to bank runs, under complete and incomplete information of the failure cluster, in order to minimize the contagion effects.

R. Elie is studying risk systemic propagation and its links with mean field games.

6.5. Backward stochastic (partial) differential equations with jumps and stochastic control with nonlinear expectation

A. Sulem, M.C. Quenez and R. Dumitrescu have studied optimization problems for BSDEs with jumps [11], optimal stopping for dynamic risk measures induced by BSDEs with jumps and associated reflected BSDEs. [24], [80], [19]. They have also investigated optimal stopping with nonlinear expectation under ambiguity, and their links with nonlinear Hamilton Jacobi Bellman variational inequalities in the Markovian case. Moreover they have obtained dynamic programming principles for mixed optimal-stopping problems with nonlinear expectations. They have also explored the links between generalized Dynkin games and double barriers reflected BSDE with jumps [56]. Stochastic control of Itô-Lévy Processes with applications to finance are studied by A. Sulem and B. Øksendal in [25], [26]. We have also contributed to the theory of BSDEs and Forward-Backward SDEs which appear as the adjoint equations associated to stochastic maximum principles, and address various issues about the relation between information and performance in non Markovian stochastic control: In particular, in the context of jump-diffusion models under partial information, A. Sulem, C. Fontana and B. Øksendal study in [20] 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.

A. Sulem, with B. Øksendal and T. Zhang has studied optimal stopping for Stochastic Partial Differential equations and associated reflected SPDEs [91], and optimal control of Forward-Backward SDEs [90].

Stochastic maximum principles for singular mean-field games are obtained in [37] with applications to optimal irreversible investments under uncertainty.

R. Dumitrescu and C. Labart have proposed a numerical approximation for Doubly Reflected BSDEs with Jumps and RCLL obstacles [35].

R. Elie studies approximate hedging prices under various risk constraints. This is done in collaboration with P. Briand, Y. Hu, A. Matoussi, B. Bouchard, L. Moreau, J.F. Chassagneux, I. Kharroubi and R. Dumitrescu.

6.6. Option Pricing

Interest rates modeling. A. Alfonsi studies an affine term structure model for interest rates that involve Wishart diffusions (with E. Palidda) [28] . Affine term structure models (Dai and Singleton, Duffie, ...) consider vector affine diffusions. Here, we extend the Linear Gaussian Model (LGM) by including some Wishart dynamics, and to get a model that could better fit the market. We have obtained a price expansion around the LGM for Caplet and Swaption prices. Also, we present a second order discretization scheme that allow to calculate exotic prices with this model.

American Options. In joint work with Aych Bouselmi, D. Lamberton studied the asymptotic behavior of the exercise boundary near maturity for American put options in exponential Lévy models [34].

He is currently working with M. Pistorius on the approximation of American options by Canadian options, which originated from the work of Peter Carr.

Barrier Options. Numerical pricing of double barrier options is investigated by A. Zanette and coauthors in [16].

6.7. Discretization of stochastic differential equations

With his PhD student A. Al Gerbi and E. Clément, B. Jourdain is interested in the strong convergence properties of the Ninomiya-Victoir scheme which is known to exhibit order 2 of weak convergence. This study is aimed at analysing the use of this scheme either at each level or only at the finest level of a multilevel Monte Carlo estimator : indeed, the variance of a multilevel Monte Carlo estimator is related to the strong error between the two schemes used in the coarse and fine grids at each level. They prove strong convergence with order $1/2$ which is improved to order 1 when the vector fields corresponding to each Brownian coordinate in the SDE commute. They also check that the renormalized errors converge to affine SDEs with source terms involving the Lie brackets between these vector fields and, in the commuting case, their Lie brackets with the drift vector field. Last, they propose a modified Ninomiya-Victoir scheme, which, at the finest level of the multilevel Monte Carlo estimator, may be coupled with strong order 1 to a simpler scheme with weak order 1 recently proposed by Giles and Szpruch.

Using optimal transport tools, A. Alfonsi, B. Jourdain and A. Kohatsu-Higa have proved that the Wasserstein distance between the time marginals of an elliptic SDE and its Euler discretization with N steps is not larger than $\frac{C\sqrt{\log(N)}}{N}$. The logarithmic factor may be removed when the uniform time-grid is replaced by a grid still counting N points but refined near the origin of times [4] . 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 using the theory developed by Ambrosio Gigli and Savare. The abstract properties of the optimal coupling between the time marginals then enable them to estimate this time derivative [30].

6.8. Advanced Monte Carlo methods.

- **Adaptive variance reduction methods.** B. Jourdain and J. Lelong have pursued their work on adaptive Monte Carlo methods in several directions [17], [36].

- **Metropolis Hastings algorithm in large dimension.** With T. Lelièvre and B. Miasojedow, B. Jourdain considers the Random Walk Metropolis algorithm on $\mathbf{R}^n$ with Gaussian proposals, and when the target probability measure is the n-fold product of a one dimensional law. It is well-known that, in the limit n tends to infinity, starting at equilibrium and for an appropriate scaling of the variance and of the timescale as a function of the dimension n, a diffusive limit is obtained for each component of the Markov chain. They generalize this result when the initial distribution is not the target probability measure ([65]). The obtained diffusive limit is the solution to a stochastic differential equation nonlinear in the sense of McKean. In [64], they prove convergence to equilibrium for this equation. They also discuss practical counterparts in order to optimize the variance of the proposal distribution to accelerate convergence to equilibrium. The analysis confirms the interest of the constant acceptance rate strategy (with acceptance rate between 1/4 and 1/3).

6.9. Numerical Probability

6.9.1. Regularity of probability laws using an interpolation method

This work was motivated by previous studies by N. Fournier, J. Printemps, E. Clément, A. Debusche and V. Bally, on the regularity of the law of the solutions of stochastic differential equations with low regularity coefficients - such as diffusion processes with Hölder coefficients or many other examples including jump type equations, Boltzmann equation or Stochastic PDE's. Since we do not have sufficient regularity, the usual approach by Malliavin calculus fails in this framework. We use the following alternative idea: 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. Consequently we are able to establish integration by parts formulas for $X(n)$, 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 its derivatives. Note 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 faster 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 knows how 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 [48] by V. Bally and Lucia Caramellino. As an application we discussed in [50] the regularity of the law of a Wiener functional under a Hörmander type non degeneracy condition.

6.9.2. A stochastic parametrix representation for the density of a Markov process.

Classical results of 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 is known as the "parametrix method". V. Bally In collaboration with A. Kohatzu Higa gave a probabilistic approach which represents the probabilistic counterpart of the parametrix method [33]. They obtained a probabilistic representation for the density of the law of the solution of a SDE and more generally, 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 - and this does not seem trivial.

6.9.3. The distance between two density functions and convergence in total variation.

V. Bally and L. Caramellino have obtained estimates of the distance between the densities of the law of two random variables using an abstract variant of Malliavin calculus. They used these estimates in order to

study the convergence in total variation of a sequence of random variables. This has been done in [49]. They are now working on more specific examples concerning the Central Limit Theorem [32]. 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 by S. Bobkov, F. Götze, G. Peccati, Y. Nourdin, D. Nualart and G. Poly. This is a very active research. Moreover, in an working paper in collaboration with his Phd student R. Clement, V. Bally uses similar methods in order to study the total variation distance between two Markov semigroups and for approximation schemes purposes. A special interest is devoted to higher order schemes such as the Victoir Nyomia scheme.

6.9.4. An invariance principle for stochastic series (U- Statistics).

Vlad Bally and Lucia Caramellino are working on invariance principles for stochastic series of polynomial type. In the case of polynomials of degree one we must have the classical Central Limit Theorem (for random variables which are not identically distributed). For polynomials of higher order we are in the framework of the so called U statistics which have been introduced by Hoffdings in t 1948 and which play an important role in modern statistics. Our contribution in this topic concerns convergence in total variation distance for this type of objects. We use abstract Malliavin calculus and more generally, the methods mentioned in the above paragraph.

MOKAPLAN Team

6. New Results

6.1. Highlights of the Year

All of the new results below are important break through and most of them non-incremental research.

Mokaplan has extended its collaborations to several researchers at Ceremade and is under review to become a project team.

6.2. Iterative Bregman Projections for Regularized Transportation Problems

Benamou, Jean-David and Carlier, Guillaume and Cuturi, Marco and Nenna, Luca and Peyré, Gabriel

[19]

We provide a general numerical framework to approximate solutions to linear programs related to optimal transport. The general idea is to introduce an entropic regularization of the initial linear program. This regularized problem corresponds to a Kullback-Leibler Bregman di-vergence projection of a vector (representing some initial joint distribution) on the polytope of constraints. We show that for many problems related to optimal transport, the set of linear constraints can be split in an intersection of a few simple constraints, for which the projections can be computed in closed form. This allows us to make use of iterative Bregman projections (when there are only equality constraints) or more generally Bregman-Dykstra iterations (when inequality constraints are involved). We illustrate the usefulness of this approach to several variational problems related to optimal transport: barycenters for the optimal trans-port metric, tomographic reconstruction, multi-marginal optimal trans-port and in particular its application to Brenier's relaxed solutions of in-compressible Euler equations, partial unbalanced optimal transport and optimal transport with capacity constraints.
The extension of the method to the Principal Agent problem, Density Functional theory and Transport under martingal constraint is under way.

6.3. A viscosity framework for computing Pogorelov solutions of the Monge-Ampere equation

Benamou, Jean-David and Froese, Brittany D.

[21]

We consider the Monge-Kantorovich optimal transportation problem between two measures, one of which is a weighted sum of Diracs. This problem is traditionally solved using expensive geometric methods. It can also be reformulated as an elliptic partial differential equation known as the Monge-Ampere equation. However, existing numerical methods for this non-linear PDE require the measures to have finite density. We introduce a new formulation that couples the viscosity and Aleksandrov solution definitions and show that it is equivalent to the original problem. Moreover, we describe a local reformulation of the subgradient measure at the Diracs, which makes use of one-sided directional derivatives. This leads to a consistent, monotone discretisation of the equation. Computational results demonstrate the correctness of this scheme when methods designed for conventional viscosity solutions fail.
The method offers a new insight into the duality between Aleksandrov and Brenier solutions of the Monge Ampère equations. We still work on the viscosity existence/uniqueness convergence of sheme theory.

6.4. Discretization of functionals involving the Monge-Ampère operator

Benamou, Jean-David and Carlier, Guillaume and Mérigot, Quentin and Oudet, Edouard

[26]

Gradient flows in the Wasserstein space have become a powerful tool in the analysis of diffusion equations, following the seminal work of Jordan, Kinderlehrer and Otto (JKO). The numerical applications of this formulation have been limited by the difficulty to compute the Wasserstein distance in dimension larger than 2. One step of the JKO scheme is equivalent to a variational problem on the space of convex functions, which involves the Monge-Ampère operator. Convexity constraints are notably difficult to handle numerically, but in our setting the internal energy plays the role of a barrier for these constraints. This enables us to introduce a consistent discretization, which inherits convexity properties of the continuous variational problem. We show the effectiveness of our approach on nonlinear diffusion and crowd-motion models.

6.5. Augmented Lagrangian methods for transport optimization, Mean-Field Games and degenerate PDEs

Benamou, Jean-David and Carlier, Guillaume

[18]

Many problems from mass transport can be reformulated as variational problems under a prescribed divergence constraint (static problems) or subject to a time dependent continuity equation which again can also be formulated as a divergence constraint but in time and space. The variational class of Mean-Field Games introduced by Lasry and Lions may also be interpreted as a generalisation of the time-dependent optimal transport problem. Following Benamou and Brenier, we show that augmented Lagrangian methods are well-suited to treat convex but nonsmooth problems. It includes in particular Monge historic optimal transport problem. A Finite Element discretization and implementation of the method is used to provide numerical simulations and a convergence study.
We have good hopes to use this method to many non-linear diffusion equations through the use of JKO gradient schemes.

6.6. Discretization of functionals involving the Monge-Ampère operator

Benamou, Jean-David and Collino, Francis and Mirebeau, Jean-Marie

[20]

We introduce a novel discretization of the Monge-Ampere operator, simultaneously consistent and degenerate elliptic, hence accurate and robust in applications. These properties are achieved by exploiting the arithmetic structure of the discrete domain, assumed to be a two dimensional cartesian grid. The construction of our scheme is simple, but its analysis relies on original tools seldom encountered in numerical analysis, such as the geometry of two dimensional lattices, and an arithmetic structure called the Stern-Brocot tree. Numerical experiments illustrate the method's efficiency.

6.7. A Γ-Convergence Result for the Upper Bound Limit Analysis of Plates

Bleyer, Jérémy and Carlier, Guillaume and Duval, Vincent and Mirebeau, Jean-Marie and Peyré, Gabriel

[23]

Upper bound limit analysis allows one to evaluate directly the ultimate load of structures without performing a cumbersome incremental analysis. In order to numerically apply this method to thin plates in bending, several authors have proposed to use various finite elements discretizations. We provide in this paper a mathematical analysis which ensures the convergence of the finite element method, even with finite elements with discontinuous derivatives such as the quadratic 6 node Lagrange triangles and the cubic Hermite triangles. More precisely, we prove the Gamma-convergence of the discretized problems towards the continuous limit analysis problem. Numerical results illustrate the relevance of this analysis for the yield design of both homogeneous and non-homogeneous materials.

6.8. Cournot-Nash equilibria

Carlier, Guillaume and Blanchet, Adrien

[24]

The notion of Nash equilibria plays a key role in the analysis of strategic interactions in the framework of N player games. Analysis of Nash equilibria is however a complex issue when the number of players is large. It is therefore natural to investigate the continuous limit as N tends to infinity and to investigate whether it corresponds to the notion of Cournot-Nash equilibria. In [9], this kind of convergence result is studied in a Wasserstein framework. In [BC1], we go one step further by giving a class of games with a continnum of players for which equilibria may be found as minimizers as a functional on measures which is very similar to the one-step JKO case, uniqueness results are the obtained from displacement convexity arguments. Finally, in [9] some situations which are non variational are considered and existence is obtained by methods combining fixed point arguments and optimal transport.

6.9. Principal Agent

Carlier, Guillaume, Benamou, Jean-David and Dupuis Xavier

The numerical resolution of principal Agent for a bilinear utility has been attacked and solved successfully in a series of recent papers see [70] and references therein.

A Bregman approach inspired by [6] has been developed for more general functions the paper is currently being written. It would be extremely useful as a complement to the theoretical analysis. A new semi-Discrete Geometric approach is also investigated where the method reduces to non-convex polynomial optimization.

6.10. Exact Support Recovery for Sparse Spikes Deconvolution

Duval, Vincent and Peyré, Gabriel

[17]

We study sparse spikes deconvolution over the space of measures. We focus our attention to the recovery properties of the support of the measure, i.e. the location of the Dirac masses. For non-degenerate sums of Diracs, we show that, when the signal-to-noise ratio is large enough, total variation regularization (which is the natural extension of the L1 norm of vectors to the setting of measures) recovers the exact same number of Diracs. We also show that both the locations and the heights of these Diracs converge toward those of the input measure when the noise drops to zero. The exact speed of convergence is governed by a specific dual certificate, which can be computed by solving a linear system. We draw connections between the support of the recovered measure on a continuous domain and on a discretized grid. We show that when the signal-to-noise level is large enough, the solution of the discretized problem is supported on pairs of Diracs which are neighbors of the Diracs of the input measure. This gives a precise description of the convergence of the solution of the discretized problem toward the solution of the continuous grid-free problem, as the grid size tends to zero.

QUANTIC Team

5. New Results

5.1. Highlights of the Year

- Experimental results in continuous measurement of error syndromes for a quantum error correction scheme developed by Mazyar Mirrahimi and his former PhD student Zaki Leghtas in close collaboration with the teams of Michel Devoret and Robert Schoelkopf (Department of Applied Physics of Yale University) have been published in Nature [13].
- Theoretical proposal on a new paradigm for universal quantum computation [12] has been chosen by the editors of the New Journal of Physics as an IOPselect paper for the novelty, significance and potential impact on future research.
- The EPOQ2 ANR Young Researcher project, led by Mazyar Mirrahimi, was highlighted in the 2013 annual report of Agence Nationale de la Recherche.

5.2. Dynamically protected cat-qubits: a new paradigm for universal quantum computation

Participant: Mazyar Mirrahimi.

In a close collaboration with the teams of Michel Devoret, Robert Schoelkopf and Liang Jiang (Department of Applied Physics, Yale university) and in particular a former member of our group, Zaki Leghtas, we have presented a new hardware-efficient paradigm for universal quantum computation. This paradigm is based on encoding, protecting and manipulating quantum information in a quantum harmonic oscillator. This proposal exploits multi-photon driven dissipative processes to encode quantum information in logical bases composed of Schrödinger cat states. More precisely, we consider two schemes. In a first scheme, a two-photon driven dissipative process is used to stabilize a logical qubit basis of two-component Schrödinger cat states. While such a scheme ensures a protection of the logical qubit against the photon dephasing errors, the prominent error channel of single-photon loss induces bit-flip type errors that cannot be corrected. Therefore, we have considered a second scheme based on a four-photon driven dissipative process which leads to the choice of four-component Schrödinger cat states as the logical qubit. Such a logical qubit can be protected against single-photon loss by continuous photon number parity measurements. Next, applying some specific Hamiltonians, we have provided a set of universal quantum gates on the encoded qubits of each of the two schemes. In particular, we have illustrated how these operations can be rendered fault-tolerant with respect to various decoherence channels of participating quantum systems. Finally, we have also proposed experimental schemes based on quantum superconducting circuits and inspired by methods used in Josephson parametric amplification, which should allow to achieve these driven dissipative processes along with the Hamiltonians ensuring the universal operations in an efficient manner.

This proposal was published in New Journal of Physics [12] and has also been chosen by the editor as an IOPselect paper for the novelty, significance and potential impact on future research.

5.3. Tracking photon jumps with repeated quantum non-demolition parity measurements

Participant: Mazyar Mirrahimi.

Quantum error correction (QEC) is required for a practical quantum computer because of the fragile nature of quantum information. In quantum error correction, information is redundantly stored in a large quantum state space and one or more observables must be monitored to reveal the occurrence of an error, without disturbing the information encoded in an unknown quantum state. Such observables, typically multi-quantum-bit parities, must correspond to a special symmetry property inherent in the encoding scheme. Measurements of these observables, or error syndromes, must also be performed in a quantum non-demolition way (projecting without further perturbing the state) and more quickly than errors occur. Previously, quantum non-demolition measurements of quantum jumps between states of well-defined energy have been performed in systems such as trapped ions, electrons, cavity quantum electrodynamics, nitrogen?vacancy centres and superconducting quantum bits. So far, however, no fast and repeated monitoring of an error syndrome had been achieved. Mazyar Mirrahimi has participated to an experiment performed by the group of Robert Schoelkopf (Department of Applied Physics, Yale University) where the quantum jumps of a possible error syndrome, namely the photon number parity of a microwave cavity, were tracked by mapping this property onto an ancilla quantum bit, whose only role is to facilitate quantum state manipulation and measurement. This quantity is just the error syndrome required in a QEC scheme proposed by Mazyar Mirrahimi and his former PhD student, Zaki Leghtas, and in a close collaboration with the teams of Michel Devoret and Robert Schoelkopf. This scheme should lead to a hardware-efficient protected quantum memory using Schrödinger cat states (quantum superpositions of different coherent states of light) in a harmonic oscillator [4]. We demonstrated the projective nature of this measurement onto a region of state space with well-defined parity by observing the collapse of a coherent state onto even or odd cat states. The measurement is fast compared with the cavity lifetime, has a high single-shot fidelity and has a 99.8 per cent probability per single measurement of leaving the parity unchanged. In combination with the deterministic encoding of quantum information in cat states realized earlier [10], the quantum non-demolition parity tracking that we have demonstrated represents an important step towards implementing an active system that extends the lifetime of a quantum bit. This result was published in Nature [9].

5.4. Dissipation-induced continuous quantum error correction for superconducting circuits

Participants: Joachim Cohen, Mazyar Mirrahimi.

Quantum error correction (QEC) is a crucial step towards long coherence times required for efficient quantum information processing (QIP). One major challenge in this direction concerns the fast real-time analysis of error syndrome measurements and the associated feedback control. Recent proposals on autonomous QEC (AQEC) have opened new perspectives to overcome this difficulty. As a sequel to our recent contributions to autonomous stabilization of maximally entangled states of superconducting qubits [53],[8], we have designed an AQEC scheme based on quantum reservoir engineering adapted to superconducting qubits. We have focused on a three-qubit bit-flip code, where three transmon qubits are dispersively coupled to a few low-Q resonator modes. By applying only continuous-wave drives of fixed but well-chosen frequencies and amplitudes, we engineer an effective interaction Hamiltonian to evacuate the entropy created by eventual bit-flip errors. We have provided a full analytical and numerical study of the protocol, while introducing the main limitations on the achievable error correction rates. This result was published in Physical Review A [11].

5.5. Continuous generation and stabilization of mesoscopic field superposition states in a quantum circuit

Participants: Ananda Roy, Mazyar Mirrahimi.

While dissipation is widely considered as being harmful for quantum coherence, it can, when properly engineered, lead to the stabilization of non-trivial pure quantum states. In a close collaboration with the teams of Michel Devoret and Douglas Stone (Department of Applied Physics, Yale University), and in the framework of a 6 months visit by Ananda Roy (PhD student at Yale), we proposed a scheme for continuous generation and stabilization of Schrödinger cat states in a cavity using dissipation engineering [15]. The scheme consists in first generating non-classical photon states with definite parity by means of a two-photon

drive and dissipation, and then stabilizing these transient states against single-photon decay. The single-photon stabilization is autonomous, and is implemented through a second engineered bath, which exploits the photon number dependent frequency-splitting due to Kerr interactions in the strongly dispersive regime of circuit QED. Starting with the Hamiltonian of the baths plus cavity, we derived an effective model of only the cavity photon states along with analytic expressions for relevant physical quantities, such as the stabilization rate. The deterministic generation of such cat states is one of the key ingredients in performing universal quantum computation.

5.6. Extending robustness and randomization from consensus to symmetrization algorithms

Participant: Alain Sarlette.

In the framework of a collaboration with Francesco Ticozzi (University of Padova) on common points between quantum and classical network dynamics, we developed a general "symmetrization" framework which covers robust ways to generate dynamics in several algorithmic and control contexts [18]. The starting point was the question of generalizing so-called "consensus" algorithms to networks composed of quantum units. In order to define state information exchange without requiring state communication (an impossible feat given the quantum no-cloning theorem), an operational viewpoint on consensus had been proposed by Alain Sarlette and co-authors in the previous year. In this new result, the scope of this operational viewpoint is considerably extended by considering it as a "symmetrization" procedure with respect to some discrete group, completely abstracting away the actual action space. It is shown that this abstraction covers existing procedures ranging from network synchronization to random state generation (not in networks) and averaging-based open-loop control procedures. The interest of viewing those procedures under the common "symmetrization" framework proposed is twofold: convergence proofs follow from a general result that we have established; and robustness to randomized actions and (specific) parameter uncertainties is shown to carry over from the "consensus" literature. It is further anticipated that the approach might be a guideline for new algorithmic designs in the future.

5.7. Accelerating consensus by spectral clustering and polynomial filters

Participant: Alain Sarlette.

The previous work of Alain Sarlette about quantum consensus and symmetrization has been further explored towards quantum-induced accelerations of algorithms, thermalization processes and random walks. This work is still at a preliminary stage. It has been noticed that some non-quantum acceleration possibilities were not fully explored and this has led to two publications that establish preliminary clarifications for our main goal. In [17], a standing conjecture has been proved which claims that if only the spectral gap of a graph is known (i.e. a bound on its lowest and largest eigenvalues), then by adding m local memories to each node no faster convergence can be obtained than by adding $m = 2$ local memories. The conjecture is proved with an analogy to root locus techniques, and a network-centric (e.g. information-theory-based) argument for this fact is currently missing, but at least the fact has been established. This allows for direct comparisons with "quantum random walk" accelerations, which obtain the same speed as $m = 2$ but with a different tweak, that is based among others on more knowledge of the network structure. In this spirit, we have clarified in [16] how classical consensus with time-varying filters can benefit from knowledge of extra bounds on the graph eigenvalue locations (without knowing them exactly, which is the case considered in the existing literature). This work also observes how the speed-up trades off with robustness to network modifications.

5.8. Integral control on Lie groups

Participant: Alain Sarlette.

A big challenge for the long-term control of interacting networks is their robustness to systematic biases. Integral control is a standard way to counter them when a target output can be measured. This method has been originally proposed, and extensively studied, for linear systems. However when the system (output) evolves on a nonlinear state space, the standard "integration" technique cannot be straightforwardly applied. Especially for global motions on spaces like the circle, sphere or (real or complex) rotation groups, the output integration viewpoint becomes problematic. We have hence proposed a new viewpoint on integral control, based on integrating the intended input [19]. For linear state spaces, it is equivalent to the standard definition. For nonlinear state spaces, this viewpoint can be transposed verbatim modulo introduction of a transport map on the tangent bundle, which is almost always present for control design purposes. In particular for systems on Lie groups, which are ubiquitous in robotics and in quantum physics, a full analysis of fully actuated systems has been proposed. The more challenging extension to underactuated systems is underway.

SIERRA Project-Team

5. New Results

5.1. An Optimal Affine Invariant Smooth Minimization Algorithmn

Participant: Alexandre d'Aspremont.

We formulate an affine invariant implementation of the algorithm in Nesterov (1983). We show that the complexity bound is then proportional to an affine invariant regularity constant defined with respect to the Minkowski gauge of the feasible set. We also detail matching lower bounds when the feasible set is an ℓp ball. In this setting, our bounds on iteration complexity for the algorithm in Nesterov (1983) are thus optimal in terms of target precision, smoothness and problem dimension. (in collaboration with Cristóbal Guzmán, Martin Jaggi)

5.2. SAGA: A Fast Incremental Gradient Method With Support for Non-Strongly Convex Composite Objectives

Participants: Simon Lacoste-Julien, Francis Bach.

In this work we introduce a new optimisation method called SAGA in the spirit of SAG, SDCA, MISO and SVRG, a set of recently proposed incremental gradient algorithms with fast linear convergence rates. SAGA improves on the theory behind SAG and SVRG, with better theoretical convergence rates, and has support for composite objectives where a proximal operator is used on the regulariser. Unlike SDCA, SAGA supports non-strongly convex problems directly, and is adaptive to any inherent strong convexity of the problem. Moreover, the proof of the convergence bounds is much simpler than the one of our earlier work SAG. (in collaboration with A. Defazio, ANU)

5.3. Non-parametric Stochastic Approximation with Large Step sizes

Participants: Aymeric Dieuleveut, Francis Bach.

We consider the random-design least-squares regression problem within the reproducing kernel Hilbert space (RKHS) framework. Given a stream of independent and identically distributed input/output data, we aim to learn a regression function within an RKHS $\mathcal{H}$, even if the optimal predictor (i.e., the conditional expectation) is not in $\mathcal{H}$. In a stochastic approximation framework where the estimator is updated after each observation, we show that the averaged unregularized least-mean-square algorithm (a form of stochastic gradient), given a sufficient large step-size, attains optimal rates of convergence for a variety of regimes for the smoothnesses of the optimal prediction function and the functions in $\mathcal{H}$.

5.4. Adaptivity of averaged stochastic gradient descent to local strong convexity for logistic regression

Participant: Francis Bach.

In this work, we consider supervised learning problems such as logistic regression and study the stochastic gradient method with averaging, in the usual stochastic approximation setting where observations are used only once. We show that after N iterations, with a constant step-size proportional to $1/R^2\sqrt{N}$ where N is the number of observations and R is the maximum norm of the observations, the convergence rate is always of order $O(1/\sqrt{N})$, and improves to $O(R^2/\mu N)$ where μ is the lowest eigenvalue of the Hessian at the global optimum (when this eigenvalue is greater than $R^2/\sqrt{N}$). Since μ does not need to be known in advance, this shows that averaged stochastic gradient is adaptive to *unknown local* strong convexity of the objective function. Our proof relies on the generalized self-concordance properties of the logistic loss and thus extends to all generalized linear models with uniformly bounded features.

5.5. Serialrank: Spectral Ranking using Seriation

Participants: Fajwel Fogel, Alexandre d'Aspremont.

We describe a seriation algorithm for ranking a set of n items given pairwise comparisons between these items. Intuitively, the algorithm assigns similar rankings to items that compare similarly with all others. It does so by constructing a similarity matrix from pairwise comparisons, using seriation methods to reorder this matrix and construct a ranking. We first show that this spectral seriation algorithm recovers the true ranking when all pairwise comparisons are observed and consistent with a total order. We then show that ranking reconstruction is still exact even when some pairwise comparisons are corrupted or missing, and that seriation based spectral ranking is more robust to noise than other scoring methods. An additional benefit of the seriation formulation is that it allows us to solve semi-supervised ranking problems. Experiments on both synthetic and real datasets demonstrate that seriation based spectral ranking achieves competitive and in some cases superior performance compared to classical ranking methods. (in coolaboration with Milan Vojnovic, Microsoft Research).

5.6. Sequential Kernel Herding: Frank-Wolfe Optimization for Particle Filtering

Participants: Simon Lacoste-Julien, Francis Bach.

Recently, the Frank-Wolfe optimization algorithm was suggested as a procedure to obtain adaptive quadrature rules for integrals of functions in a reproducing kernel Hilbert space (RKHS) with a potentially faster rate of convergence than Monte Carlo integration (and "kernel herding" was shown to be a special case of this procedure). In this paper, we propose to replace the random sampling step in a particle filter by Frank-Wolfe optimization. By optimizing the position of the particles, we can obtain better accuracy than random or quasi-Monte Carlo sampling. In applications where the evaluation of the emission probabilities is expensive (such as in robot localization), the additional computational cost to generate the particles through optimization can be justified. Experiments on standard synthetic examples as well as on a robot localization task indicate indeed an improvement of accuracy over random and quasi-Monte Carlo sampling. (in collaboration with Fredrik Lindsten, Cambridge University)

5.7. Learning to Learn for Structured Sparsity

Participants: Nino Shervashidze, Francis Bach.

Structured sparsity has recently emerged in statistics, machine learning and signal processing as a promising paradigm for learning in high-dimensional settings. All existing methods for learning under the assumption of structured sparsity rely on prior knowledge on how to weight (or how to penalize) individual subsets of variables during the subset selection process, which is not available in general. Inferring group weights from data is a key open research problem in structured sparsity.

In this work, we propose a Bayesian approach to the problem of group weight learning. We model the group weights as hyperparameters of heavy-tailed priors on groups of variables and derive an approximate inference scheme to infer these hyperparameters. We empirically show that we are able to recover the model hyperparameters when the data are generated from the model, and we demonstrate the utility of learning weights in synthetic and real denoising problems.

5.8. Analysis of purely random forests bias

Participant: Sylvain Arlot.

Random forests are a very effective and commonly used statistical method, but their full theoretical analysis is still an open problem. As a first step, simplified models such as purely random forests have been introduced, in order to shed light on the good performance of random forests. In this paper, we study the approximation error (the bias) of some purely random forest models in a regression framework, focusing in particular on the influence of the number of trees in the forest. Under some regularity assumptions on the regression function, we show that the bias of an infinite forest decreases at a faster rate (with respect to the size of each tree) than a single tree. As a consequence, infinite forests attain a strictly better risk rate (with respect to the sample size) than single trees. Furthermore, our results allow to derive a minimum number of trees sufficient to reach the same rate as an infinite forest. As a by-product of our analysis, we also show a link between the bias of purely random forests and the bias of some kernel estimators. (In collaboration with Robin Genuer, Université de Bordeaux)

5.9. Large-Margin Metric Learning for Constrained Partitioning Problems

Participants: Rémi Lajugie, Sylvain Arlot, Francis Bach.

We consider unsupervised partitioning problems based explicitly or implicitly on the minimization of Euclidean distortions, such as clustering, image or video segmentation, and other change-point detection problems. We emphasize on cases with specific structure, which include many practical situations ranging from mean-based change-point detection to image segmentation problems. We aim at learning 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 (partially) labeled 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. Our experiments show how learning the metric can significantly improve performance on bioinformatics, video or image segmentation problems.

5.10. Metric Learning for Aligning temporal sequences

Participants: Damien Garreau, Rémi Lajugie, Sylvain Arlot, Francis Bach.

In this work, we propose to learn a Mahalanobis distance to perform alignment of multivariate time series. The learning examples for this task are time series for which the true alignment is known. We cast the alignment problem as a structured prediction task, and propose realistic losses between alignments for which the optimization is tractable. We provide experiments on real data in the audio to audio context, where we show that the learning of a similarity measure leads to improvements in the performance of the alignment task. We also propose to use this metric learning framework to perform feature selection and, from basic audio features, build a combination of these with better performance for the alignment.

5.11. Weakly Supervised Action Labeling in Videos Under Ordering Constraints

Participants: Rémi Lajugie, Francis Bach.

We are given a set of video clips, each one annotated with an ordered list of actions, such as “walk” then “sit” then “answer phone” extracted from, for example, the associated text script. We seek to temporally localize the individual actions in each clip as well as to learn a discriminative classifier for each action. We formulate the problem as a weakly supervised temporal assignment with ordering constraints. Each video clip is divided into small time intervals and each time interval of each video clip is assigned one action label, while respecting the order in which the action labels appear in the given annotations. We show that the action label assignment can be determined together with learning a classifier for each action in a discriminative manner. We evaluate the proposed model on a new and challenging dataset of 937 video clips with a total of 787720 frames containing sequences of 16 different actions from 69 Hollywood movies. (in collaboration with Piotr Bojanowski, Ivan Laptev, Jean Ponce, Cordelia Schmid and Josef Sivic)

5.12. On Pairwise Cost for Multi-Object Network Flow Tracking

Participant: Simon Lacoste-Julien.

Multi-object tracking has been recently approached with the min-cost network flow optimization techniques. Such methods simultaneously resolve multiple object tracks in a video and enable modeling of dependencies among tracks. Min-cost network flow methods also fit well within the "tracking-by-detection" paradigm where object trajectories are obtained by connecting per-frame outputs of an object detector. Object detectors, however, often fail due to occlusions and clutter in the video. To cope with such situations, we propose an approach that regularizes the tracker by adding second order costs to the min-cost network flow framework. While solving such a problem with integer variables is NP-hard, we present a convex relaxation with an efficient rounding heuristic which empirically gives certificates of small suboptimality. Results are shown on real-world video sequences and demonstrate that the new constraints help selecting longer and more accurate tracks improving over the baseline tracking-by-detection method. (in collaboration with Visesh Chari,Ivan Laptev, Josef Sivic).

5.13. A Markovian approach to distributional semantics with application to semantic compositionality

Participants: Edouard Grave, Francis Bach, Guillaume Obozinski.

In this work, we describe a new approach to distributional semantics. This approach relies on a generative model of sentences with latent variables, which takes the syntax into account by using syntactic dependency trees. Words are then represented as posterior distributions over those latent classes, and the model allows to naturally obtain in-context and out-of-context word representations, which are comparable. We train our model on a large corpus and demonstrate the compositionality capabilities of our approach on different datasets.

5.14. A convex relaxation for weakly supervised relation extraction

Participant: Edouard Grave.

A promising approach to relation extraction, called weak or distant supervision, exploits an existing database of facts as training data, by aligning it to an unlabeled collection of text documents. Using this approach, the task of relation extraction can easily be scaled to hundreds of different relationships. However, distant supervision leads to a challenging multiple instance, multiple label learning problem. Most of the proposed solutions to this problem are based on non-convex formulations, and are thus prone to local minima. In this article, we propose a new approach to the problem of weakly supervised relation extraction, based on discriminative clustering and leading to a convex formulation. We demonstrate that our approach outperforms state-of-the-art methods on a challenging dataset introduced in 2010.

5.15. Weakly supervised named entity classification

Participant: Edouard Grave.

In this paper, we describe a new method for the problem of named entity classification for specialized or technical domains, using distant supervision. Our approach relies on a simple observation: in some specialized domains, named entities are almost unambiguous. Thus, given a seed list of names of entities, it is cheap and easy to obtain positive examples from unlabeled texts using a simple string match. Those positive examples can then be used to train a named entity classifier, by using the PU learning paradigm, which is learning from positive and unlabeled examples. We introduce a new convex formulation to solve this problem, and apply our technique in order to extract named entities from financial reports corresponding to healthcare companies.

5.16. Fast imbalanced binary classification: a moment-based approach

Participant: Edouard Grave.

In this paper, we consider the problem of imbalanced binary classification in which the number of negative examples is much larger than the number of positive examples. The two mainstream methods to deal with such problems are to assign different weights to negative and positive points or to subsample points from the negative class. In this paper, we propose a different approach: we represent the negative class by the two first moments of its probability distribution (the mean and the covariance), while still modeling the positive class by individual examples. Therefore, our formulation does not depend on the number of negative examples, making it suitable to highly imbalanced problems and scalable to large datasets. We demonstrate empirically, on a protein classification task and a text classification task, that our approach achieves similar statistical performance than the two mainstream approaches to imbalanced classification problems, while being more computationally efficient. (in collaboration with Laurent El Ghaoui, U.C. Berkeley)

ANGE Project-Team

6. New Results

6.1. Highlights of the Year

In 2014, ANGE status turned from Inria team to Inria project-team. Afterwards, M. Parisot was recruited by Inria as a junior researcher.

6.2. Analysis of models in fluid mechanics

6.2.1. Well-posedness of multilayer Shallow Water-type equations

Participants: Emmanuel Audusse, Bernard Di Martino, Ethem Nayir, Yohan Penel.

The hyperbolicity of some 2-layer Shallow Water equations had been proven in [26], [23], there are many open theoretical investigations to lead about these systems. In particular, E. Nayir proved the local well-posedness of the model derived in [23] for periodic boundary conditions. Next steps will consist in extending this preliminary result to the whole space and proving the global existence of strong solutions. The existence of weak solutions will be studied from B. Di Martino's work. The hyperbolicity for N layers must also be investigated.

As for numerical aspects, the use of FRESHKISS3D will provide qualitative assessments for modelling issues (viscous tensor, source terms, variable density, interfacial velocities). It will also yield comparisons with theoretical results, in particular when the number of layers goes to infinity.

6.2.2. Non-hydrostatic models

Participants: Dena Kazerani, Jacques Sainte-Marie, Nicolas Seguin.

Together with Corentin Audiard from Univ. Pierre et Marie Curie, we investigated the structure of general non hydrostatic models for shallow water flows. This includes the Green–Naghdi equations and the model proposed by Bristeau *et al.* in [13]. D. Kazerani proved that such systems possess a symmetric structure based on the existence of an energy. The main difference with the well-known hyperbolic case is due to the presence of differential operators instead of matrices.

6.3. Modelling of complex flows

6.3.1. Dynamics of sedimentary river beds with stochastic fluctuations

Participants: Emmanuel Audusse, Philippe Ung.

We studied in [9] the behaviour of the solution of the Saint-Venant–Exner equations when a stochastic term is introduced in the model through the sediment flux. A first investigation was done considering periodic boundary conditions and the next part of this study is devoted to the case when physical ones are imposed. Our goal is to investigate the possibility to bring out a characteristic long time behaviour and to establish a relation between the injected noise and the physical parameters involved in the model. This work was achieved in collaboration with Sébastien Boyaval from Lab. Hydraulique Saint-Venant.

6.3.2. Non-hydrostatic effects

Participants: Nora Aïssiouene, Marie-Odile Bristeau, Edwige Godlewski, 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 and to develop a robust numerical method for the resolution of the model.

Concerning the modelling aspect, a non-hydrostatic shallow water-type model approximating the incompressible Euler and Navier-Stokes sytems with free surface was developped and published in [13]. The closure relations are obtained by a minimal energy constraint instead of an asymptotic expansion. The model slightly differs from the well-known Green-Naghdi model and is confronted with stationary and analytical solutions of the Euler system corresponding to rotational flows.

The numerical approximation relies on a projection-correction type scheme. The hyperbolic part of the system is approximated using a kinetic finite volume solver and the correction step implies to solve an elliptic problem involving the non-hydrostatic part of the pressure.

In one dimension, the resolution of the incompressibility problem leads to solve a mixed problem where the pressure and the velocity are defined in compatible approximation spaces. This step uses a variationnal formulation of the shallow water version of the incompressibility condition.

This numerical scheme satisfies classical properties (positivity, well-balancing and consistency) and a discrete entropy inequality. Several numerical experiments are performed to confirm the relevance of our approach.

This approach will allow us to extend the numerical method in higher dimensions and to treat particular difficult cases occuring in specific geophysical situations (dry/wet interfaces).

6.3.3. Plasticity in Shallow Water equations

Participant: Nicolas Seguin.

In collaboration with Bruno Després and Clément Mifsud from Univ. Pierre et Marie Curie, we proposed in [20] a new definition of solutions for hyperbolic Friedrichs' systems in bounded domains, which follows the idea of Lions' dissipative solutions and Otto's boundary formulation for conservation laws. We proved in the classical settings existence and uniqueness. The goal of this project is to be able to incorporate nonlinear effects of plasticity in models of elasticity or overflowing in channels for shallow water flows, by adding entropy compatible constraints.

6.3.4. Management of marine energies

Participants: Cindy Guichard, Martin Parisot, Jacques Sainte-Marie, Julien Salomon.

The purpose of this project is to model floating devices (like buoys) in the context of recovering energy from water resources (seas and oceans). If the free surface flow can be handled by means of the Saint-Venant equations, the area under the buoys requires a different modelling (for example equivalence with springs) as the surface is constrained. The Archimedes' principle is also involved. Some preliminary numerical results were obtained thanks to the FreshKiss3D code.

To go further, the optimisation of the overall process is also under consideration. Indeed, to maximise the amount of recovered energy, the bathymetry, the shape of the buoy, the number of buoys are critical parameters which must be modelled in view of industrial applications. Optimal control methods are applied to determine the best configuration depending on the devices: optimisation of the kinetic energy for water-turbines or of the potential energy for buoys.

6.4. Accurate simulations of fluid flows

6.4.1. A numerical scheme for the Saint-Venant–Exner equations

Participants: Emmanuel Audusse, Philippe Ung.

After having established a Godunov-type method based on the design of a three-wave Approximate Riemann Solver for the Saint-Venant equations [10], we extended this approach to the Saint-Venant–Exner equations for modelling the sediment transport. The coupled aspect between the hydraulic and the morphodynamic parts is only located on the evaluation of the wave velocities. Under this assumption, the proposed scheme can be interpreted as a hybrid method between the splitting and non-splitting methods and it also raises the issue of the choice between the two previous approaches.

These results were proven in collaboration with Christophe Chalons from Univ. Versailles–Saint-Quentin.

6.4.2. Simulations of fluid/particules interactions

Participant: Nicolas Seguin.

In collaboration with Nina Aguillon and Frédéric Lagoutière from Univ. Paris-Sud, we proved in [7] the convergence of finite volume schemes for a simplified model of fluid-particle interaction. The mesh follows the particle which appears in the model as a pointwise contribution. The numerical scheme is based on local well-balanced fluxes, which permits to obtain compactness and convergence.

6.4.3. Hydrostatic reconstruction

Participants: Emmanuel Audusse, Marie-Odile Bristeau, Jacques Sainte-Marie.

The hydrostatic reconstruction is a general and efficient method to handle source terms that uses an arbitrary solver for the homogeneous problem and leads to a consistent, well-balanced, positive scheme satisfying a semi-discrete entropy inequality.

In [8], we proved with Francois Bouchut from Univ. Marne-la-Vallée that the hydrostatic reconstruction coupled to the classical kinetic solver satisfies a fully discrete entropy inequality which involves an error term but the latter goes to zero strongly with the mesh size.

6.4.4. A numerical scheme for multilayer shallow-water model for all Froude regimes

Participant: Martin Parisot.

The aim of this work in collaboration with Jean-Paul Vila from INSA/IMT is to propose an efficient numerical resolution to simulate stratified non-miscible fluids. The strategy should be consistent for all regime especially with the so-called low-Froude regime particularly relevant for applications. The proposed scheme is entropy-satisfying, well-balanced and asymptotic preserving. In addition the stability of the scheme is ensured for large time scale. More precisely, it does not depend on the gravity waves, which are very restrictive for the targeted applications, such as oceanology and meteorology. Further work using the strategy for sustainable energies is in progress.

6.4.5. Adaptation of the Godunov scheme to the low Froude regime

Participants: Emmanuel Audusse, Do Minh Hieu, Yohan Penel.

Standard numerical schemes designed for the simulation of fluid flows are known to fail when the Mach number becomes too small. Similar behaviours are observed for geophysical flows when the Froude number decreases. Do Minh Hieu is interested in the numerical simulation of the Shallow Water equations including some Coriolis forces. He investigated several corrections of the standard Godunov schemes in 1D to preserve the kernel of spatial operators involved in the aformentioned equations and blamed for being responsible of the loss of accuracy. He now intends to perform the same analysis in 2D under the supervision of E. Audusse, S. Dellacherie (from CEA), P. Omnès (from CEA) and Y. Penel.

6.5. Software development and assessments

6.5.1. Improvements in the FRESHKISS3D code

Participants: Marie-Odile Bristeau, David Froger, Raouf Hamouda, Jacques Sainte-Marie.

Several tasks have been achieved in the FRESHKISS3D software:

- FreshKiss3D has been improved to take into account the second order in space for the 3D cases.
- The solver now includes the second order in time.
- The numerical validation using 3D numerical analytical solutions has been achieved.
- Numerous simulations have been driven by industrial contracts:
  - Simulations of fluid hydrodynamics in lagoons for optimizing the geometric field to ensure a high level of agitation for a low energy consumption (SAUR)
  - Simulations of fluid hydrodynamics in lagoons showing the vertical distribution of velocity and how to use it for optimizing micro-algae production (Salinalgue)
- Tsunamis simulations leading to the module TsunaMaths, web interface showing some historical tsunamis.
- Geometric implementations of FRESHKISS3D have been improved.
- Unit tests are being made automatically as the source code is modified.
- A user interface has been created using Python.
- The parallelization of FRESHKISS3D with MPI is under development.

ARAMIS Project-Team

6. New Results

6.1. Highlights of the Year

ARAMIS has contributed to the special issue on "Complex network theory and the brain" in the prestigious journal of Philosopical Transactions of the Royal Society, Series B. This work was featured by the ICM (http://icm-institute.org/en/news/complex-network-theory-and-the-brain?lang=en) and Inria (http://www.inria.fr/en/centre/paris-rocquencourt/news/complex-network-theory-and-the-brain).

6.2. Detection of volume loss in hippocampal layers in Alzheimer's disease using 7 T MRI

Participants: Claire Boutet, Marie Chupin, Stéphane Lehéricy, Linda Marrakchi-Kacem, Stéphane Epelbaum, Cyril Poupon, Christopher Wiggins, Alexandre Vignaud, Dominique Hasboun, Bénédicte Desfontaines, Olivier Hanon, Bruno Dubois, Marie Sarazin, Lucie Hertz-Pannier, Olivier Colliot [Correspondant].

In Alzheimer's disease (AD), the hippocampus is an early site of tau pathology and neurodegeneration. Histological studies have shown that lesions are not uniformly distributed within the hippocampus. Moreover, alterations of different hippocampal layers may reflect distinct pathological processes. 7 T MRI dramatically improves the visualization of hippocampal subregions and layers. In this study, we aimed to assess whether 7 T MRI can detect volumetric changes in hippocampal layers in vivo in patients with AD. We studied four AD patients and seven control subjects. MR images were acquired using a whole-body 7 T scanner with an eight channel transmit-receive coil. Hippocampal subregions were manually segmented from coronal T2*-weighted gradient echo images with 0.3 × 0.3 × 1.2 mm3 resolution using a protocol that distinguishes between layers richer or poorer in neuronal bodies (Figure 1). Five subregions were segmented in the region of the hippocampal body: alveus, strata radiatum, lacunosum and moleculare (SRLM) of the cornu Ammonis (CA), hilum, stratum pyramidale of CA and stratum pyramidale of the subiculum. We found strong bilateral reductions in the SRLM of the cornu Ammonis and in the stratum pyramidale of the subiculum ($p < 0.05$), with average cross-sectional area reductions ranging from -29% to -49%. These results show that it is possible to detect volume loss in distinct hippocampal layers using segmentation of 7 T MRI. 7 T MRI-based segmentation is a promising tool for AD research.

More details in [3].

6.3. White matter lesions in patients with frontotemporal lobar degeneration due to progranulin mutations

Participants: Paola Caroppo, Isabelle Le Ber, Agnès Camuzat, Fabienne Clot, Lionel Naccache, Foudil Lamari, Anne Bertrand, Serge Belliard, Olivier Colliot [Correspondant], Alexis Brice.

Mutations in the progranulin (GRN) gene are responsible for 20% of familial cases of frontotemporal dementias. All cause haploinsufficiency of progranulin, a protein involved in inflammation, tissue repair, and cancer. Carriers of the GRN mutation are characterized by a variable degree of asymmetric brain atrophy, predominantly in the frontal, temporal, and parietal lobes. We described four GRN mutation carriers with remarkable widespread white matter lesions (WML) associated with lobar atrophy shown on magnetic resonance imaging. The WML were predominantly in the frontal and parietal lobes and were mostly confluent, affecting the periventricular subcortical white matter and U-fibers. In all patients, common vascular, metabolic, inflammatory, dysimmune, and mitochondrial disorders were excluded and none had severe vascular risk factors. Our data suggest that white matter involvement may be linked to progranulin pathological processes in a subset of GRN mutation carriers. The plasma progranulin measurement, which is predictive of GRN mutations, and GRN sequencing should thus be included in investigations of patients with frontotemporal lobar degenerations who show unusual white matter hyperintensities and atrophy on magnetic resonance imaging.

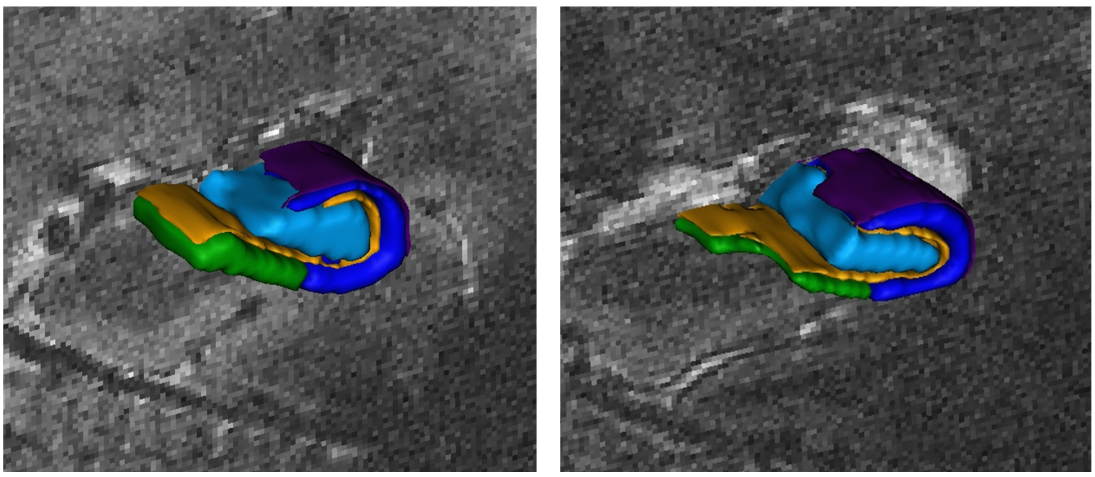


Figure 1. Segmentation of hippocampal layers using in vivo 7 Tesla MRI. were performed on the second echo image. Left panel: control subject. Right panel: patient with Alzheimer's disease. Purple, alveus; dark blue, stratum pyramidale of CA1-3; yellow, strata radiatum,lacunosum and moleculare of CA1-3, strata lacunosum and moleculare of the subiculum and stratum moleculare of gyrus dentatus; cyan, stratum pyramidale of CA4 and stratum granulosum and polymorphic layer of gyrus dentatus; green, stratum pyramidale of the subiculum.

More details in [4].

6.4. Template-based morphometry using diffeomorphic iterative centroids

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

A common approach for the analysis of anatomical variability relies on the estimation of a representative template of the population, followed by the study of this population based on the parameters of the deformations going from the template to the population. The Large Deformation Diffeomorphic Metric Mapping framework is widely used for shape analysis of anatomical structures, but computing a template with such framework is computationally expensive. We proposed a fast approach for template-based analysis of anatomical variability. The template is estimated using an iterative approach which quickly provides a centroid of the population. Statistical analysis is then performed using principal component analysis on the initial momenta that define the deformations between the centroid and each subject of the population. This approach was applied to the analysis of hippocampal shape on 80 patients with Alzheimer's Disease and 138 controls from the ADNI database.

More details in [22] and [36].

6.5. Structural connectivity differences in left and right temporal lobe epilepsy

Participants: Pierre Besson, Vera Dinkelacker [Correspondant], Romain Valabrègue, Lionel Thivard, Xavier Leclerc, Michel Baulac, Daniela Sammler, Olivier Colliot, Stéphane Lehéricy, Séverine Samson, Sophie Dupont.

Our knowledge on temporal lobe epilepsy (TLE) with hippocampal sclerosis has evolved towards the view that this syndrome affects widespread brain networks. Diffusion weighted imaging studies have shown alterations of large white matter tracts, most notably in left temporal lobe epilepsy, but the degree of altered connections between cortical and subcortical structures remains to be clarified. We performed a whole brain connectome analysis in 39 patients with refractory temporal lobe epilepsy and unilateral hippocampal sclerosis (20 right and 19 left) and 28 healthy subjects. We performed whole-brain probabilistic fiber tracking using MRtrix and segmented 164 cortical and subcortical structures with Freesurfer. Individual structural connectivity graphs based on these 164 nodes were computed by mapping the mean fractional anisotropy (FA) onto each tract. Connectomes were then compared using two complementary methods: permutation tests for pair-wise connections and Network Based Statistics to probe for differences in large network components. Comparison of pair-wise connections revealed a marked reduction of connectivity between left TLE patients and controls, which was strongly lateralized to the ipsilateral temporal lobe. Specifically, infero-lateral cortex and temporal pole were strongly affected, and so was the perisylvian cortex. In contrast, for right TLE, focal connectivity loss was much less pronounced and restricted to bilateral limbic structures and right temporal cortex. Analysis of large network components revealed furthermore that both left and right hippocampal sclerosis affected diffuse global and interhemispheric connectivity. Thus, left temporal lobe epilepsy was associated with a much more pronounced pattern of reduced FA, that included major landmarks of perisylvian language circuitry. These distinct patterns of connectivity associated with unilateral hippocampal sclerosis show how a focal pathology influences global network architecture, and how left or right-sided lesions may have differential and specific impacts on cerebral connectivity.

More details in [2].

6.6. Morphometry of anatomical shape complexes with dense deformations and sparse parameters

Participants: Stanley Durrleman [Correspondant], Marcel Prastawa, Nicolas Charon, Julie Korenberg, Sarang Joshi, Guido Gerig, Alain Trouvé.

We propose a generic method for the statistical analysis of collections of anatomical shape complexes, namely sets of surfaces that were previously segmented and labeled in a group of subjects. The method estimates an anatomical model, the template complex, that is representative of the population under study. Its shape reflects anatomical invariants within the dataset. In addition, the method automatically places control points near the most variable parts of the template complex. Vectors attached to these points are parameters of deformations of the ambient 3D space. These deformations warp the template to each subject's complex in a way that preserves the organization of the anatomical structures. Multivariate statistical analysis is applied to these deformation parameters to test for group differences. Results of the statistical analysis are then expressed in terms of deformation patterns of the template complex, and can be visualized and interpreted. The user needs only to specify the topology of the template complex and the number of control points. The method then automatically estimates the shape of the template complex, the optimal position of control points and deformation parameters. The proposed approach is completely generic with respect to any type of application and well adapted to efficient use in clinical studies, in that it does not require point correspondence across surfaces and is robust to mesh imperfections such as holes, spikes, inconsistent orientation or irregular meshing.

The approach is illustrated with a neuroimaging study of Down syndrome (DS). Results demonstrate that the complex of deep brain structures shows a statistically significant shape difference between control and DS subjects. The deformation-based modeling is able to classify subjects with very high specificity and sensitivity, thus showing important generalization capability even given a low sample size. We show that results remain significant even if the number of control points, and hence the dimension of variables in the statistical model, are drastically reduced. The analysis may even suggest that parsimonious models have an increased statistical performance.

The method has been implemented in the software Deformetrica, which is publicly available at www.deformetrica.org

More details in [14].

6.7. Iconic-Geometric Nonlinear Registration of a Basal Ganglia Atlas for Deep Brain Stimulation Planning

Participants: Ana Fouquier, Stanley Durrleman, Jérôme Yelnik, Sara Fernandez-Vidal, Eric Bardinet.

We evaluated a nonlinear registration method for warping a 3D histological atlas of the basal ganglia into patient data for deep brain stimulation (DBS) planning. The power of the method is the possibility to combine iconic registration with geometric constraints under a unified diffeomorphic framework. This combination aims to ensure robust and accurate atlas-to-patient warping and anatomy-preserving deformations of stimulation target nuclei. A comparison of the method with a state-of-the-art diffeomorphic registration algorithm reveals how each approach deforms low-contrasted image regions where DBS target nuclei often lie. The technique is applied to T1-weighted magnetic resonance images from a cohort of Parkinsonian subjects, including subjects with standard-size and large ventricles. Results illustrate the effects of iconic or geometric registration alone, as well as how both constraints can be integrated in order to contribute for registration precision enhancement. See Fig. 2 .

More details in [25].

6.8. Evaluation of morphometric descriptors of deep brain structures for the automatic classification of patients with Alzheimer's disease, mild cognitive impairment and elderly controls

Participants: Alexandre Routier [correspondant], Pietro Gori, Ana Fouquier, Sophie Lecomte, Olivier Colliot, Stanley Durrleman.

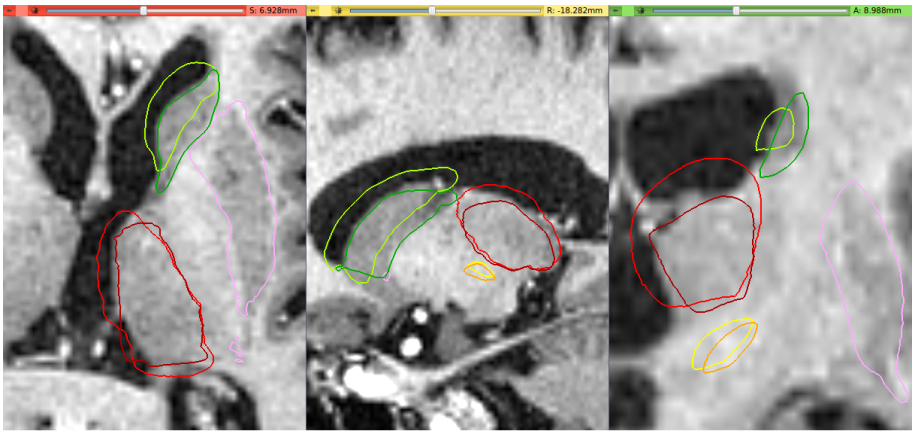


Figure 2. Superimposition of deformed meshes of the histological atlas with a patient pre-operative MRI. Meshes in bright colors result from a block-matching algorithm based on image intensity. Meshes in dark colors result from our iconic-geometric approach with non-linear deformation. We observe a better alignment of the structures, as well as a realistic deformation of the sub-thalamic nucleus (in yellow/orange), which is not visible in the image and therefore has not been taken into account for estimating the optimal deformation. This nucleus is the stimulation target for patients with Parkinson's disease.

We participated in the Computer-Aided Diagnosis of Dementia based on structural MRI data (http://caddementia.grand-challenge.org/). Our approach was to select shapes of 12 brain structures: the caudate nucleus, putamen, pallidum, thalamus, hippocampus and amygdala of each hemisphere. The structure segmentation was based on a FreeSurfer segmentation and the marching-cubes algorithm was used to get 3D triangular meshes. Using our software Deformetrica, anatomical models (mean shape and typical variations) of these brain structures were built for patients with Alzheimer's disease (AD), Mild Cognitive Impairments (MCI) and cognitively normal controls (CN) based on the data of 509 ADNI subjects. The models for AD, MCI and CN were registered to the test subjects by maximizing the likelihood of the test image to be derived from each model. The final classification was made by thresholding this criterion taking into account the covariance of the deformation parameters. The thresholds were either optimized on the ADNI data or on the provided training data. The method was fully automatic and the computation time was 4 days for training the anatomical models plus 11 hours per subject for registration and classification. For the 30 training subjects, the algorithm had accuracies of 73% (if optimized on training data) and 50% (if optimized on ADNI data). On the test set of 354 images, our method yields an accuracy of 49.2% (43.5 - 54.2), true positive fraction of 94.6% (89.8 - 97.7) for CN, 11.5% (6.2 - 17.7) for MCI and 36.9% (27.4 - 46.5) for AD.

Our participation to this challenge was the opportunity to test our software Deformetrica for classification tasks. It ran on more than 800 images, thus showing its ability to deal with large data sets.

More details in [27].

6.9. A Prototype Representation to Approximate White Matter Bundles with Weighted Currents

Participants: Pietro Gori [correspondant], Olivier Colliot, Linda Marrakchi-Kacem, Fabrizio de Vico Fallani, Mario Chavez, Sophie Lecomte, Cyril Poupon, Andreas Hartmann, Nicholas Ayache, Stanley Durrleman.

Quantitative and qualitative analysis of white matter fibers resulting from tractography algorithms is made difficult by their huge number. To this end, we propose an approximation scheme which gives as result a more concise but at the same time exhaustive representation of a fiber bundle. It is based on a novel computational model for fibers, called weighted currents, characterized by a metric that considers both the pathway and the anatomical locations of the endpoints of the fibers. Similarity has therefore a twofold connotation: geometrical and related to the connectivity. The core idea is to use this metric for approximating a fiber bundle with a set of weighted prototypes, chosen among the fibers, which represent ensembles of similar fibers. The weights are related to the number of fibers represented by the prototypes. The algorithm is divided into two steps. First, the main modes of the fiber bundle are detected using a modularity based clustering algorithm. Second, a prototype fiber selection process is carried on in each cluster separately. This permits to explain the main patterns of the fiber bundle in a fast and accurate way. See Fig. 3

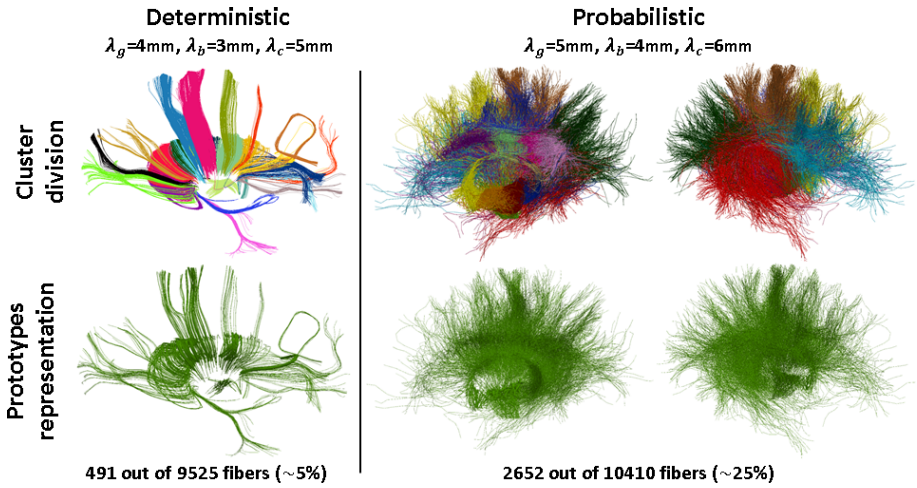


Figure 3. Illustration of our method to cluster fibers and approximate clusters based on a weighted currents metric, which measures differences in the locations of fibers extremities and the geometry of their pathway. 2 examples are shown using fibers from a deterministic tractography (left) and probabilistic tractography (right). Clustering (top row) and approximation of fibers within each cluster (bottom row) are shown.

More details in [24].

6.10. Non-parametric resampling of random walks for spectral network 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 community detection.

More details in [10].

6.11. Graph analysis of functional brain networks: practical issues in translational neurosciences

Participants: Fabrizio de Vico Fallani [correspondant], Sophie Achard, Jonas Richiardi, Mario Chavez.

The brain can be regarded as a network: a connected system where nodes, or units, represent different specialized regions and links, or connections, represent communication pathways. From a functional perspective, communication is coded by temporal dependence between the activities of different brain areas. In the last decade, the abstract representation of the brain as a graph has allowed to visualize functional brain networks and describe their non-trivial topological properties in a compact and objective way. Nowadays, the use of graph analysis in translational neuroscience has become essential to quantify brain dysfunctions in terms of aberrant reconfiguration of functional brain networks. Despite its evident impact, graph analysis of functional brain networks is not a simple toolbox that can be blindly applied to brain signals. On the one hand, it requires the know-how of all the methodological steps of the pipeline that manipulate the input brain signals and extract the functional network properties. On the other hand, knowledge of the neural phenomenon under study is required to perform physiologically relevant analysis. The aim of our work is to provide practical indications to make sense of brain network analysis and contrast counterproductive attitudes.

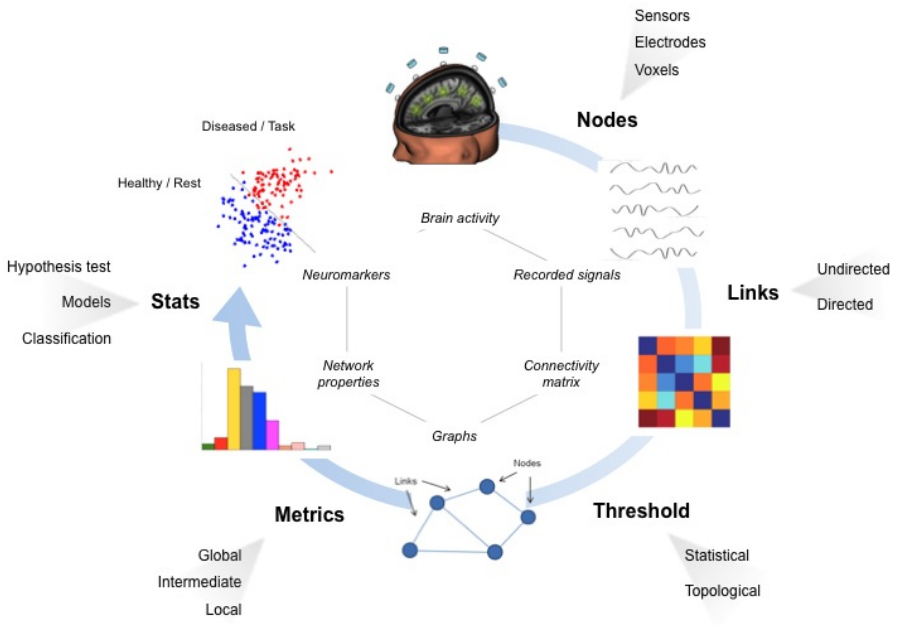


Figure 4. Processing pipeline for functional brain connectivity modeling and analysis. Nodes correspond to specific brain sites according to the used neuroimaging technique. Links are estimated by measuring the functional connectivity (FC) between the activity of brain nodes; this information is contained in a connectivity matrix. By means of filtering procedures, based on thresholds, only the most important links constitute the brain graph. The topology of the brain graph is quantified by different graph metrics (or indices) that can be represented as numbers (e.g. the colored bars). These graph indices can be input to statistical analysis to look for significant differences between populations/conditions (e.g. red points correspond to brain graph indices of diseased patients or tasks, blue points stand for healthy subjects or resting states).

More details in [11].

6.12. Hierarchy of neural organisation in the zebra fish spinal cord: causality analysis of in-vivo calcium imaging data

Participants: Fabrizio de Vico Fallani [correspondant], Martina Corazzol, Jnena Sternberg, Kevin Fidelin, Claire Wyart, Mario Chavez.

The recent development of genetically encoded calcium indicators enables monitoring in vivo the activity of neuronal populations. Most analysis of these calcium transients relies on linear regression analysis based on the sensory stimulus applied or the behavior observed. To estimate the basic properties of the functional neural circuitry, we propose a network-based approach based on calcium imaging recorded at single cell resolution. Differently from previous analysis based on cross-correlation, we used Granger causality estimates to infer activity propagation between the activities of different neurons. The resulting functional networks were then modeled as directed graphs and characterized in terms of connectivity and node centralities. We applied our approach to calcium transients recorded at low frequency (4 Hz) in ventral neurons of the zebrafish spinal cord at the embryonic stage when spontaneous coiling of the tail occurs. Our analysis on population calcium imaging data revealed a strong ipsilateral connectivity and a characteristic hierarchical organization of the network hubs that supported established propagation of activity from rostral to caudal spinal cord. Our method could be used for detecting functional defects in neuronal circuitry during development and pathological conditions.

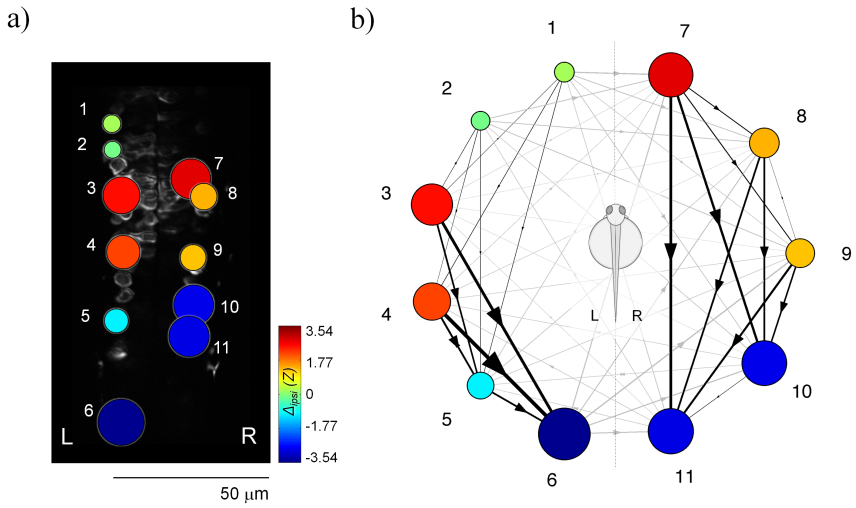


Figure 5. Rostro-caudal distribution of the nodal delta centrality in the representative zebrafish embryo. Panel a) The normalized ipsi value is represented for each node (motoneuron) as a colored circle superimposed on the field of view. The larger the circle, the more central is the node in terms of its tendency to act as a transmitter (red color, positive value) or receiver (blue color, negative value) hub of information flow. Panel b) The same normalized ipsi centrality values are here represented within the neuronal GC network. Statistically significant GC influences are illustrated as directed arrows. The thicker the arrow the stronger the GC value is. Inter-hemicord directed links are illustrated in gray color for the sake of simplicity.

More details in [9].

6.13. 2D harmonic filtering of MR phase images in multicenter clinical setting: towards a magnetic signature of cerebral microbleeds

Participants: Takoua Kaaouana [correspondant], Ludovic de Rochefort, Thomas Samaille, Nathalie Thiery, Carole Dufouil, Christine Delmaire, Didier Dormont, Marie Chupin.

Cerebral microbleeds (CMBs) have emerged as a new imaging marker of small vessel disease. Composed of hemosiderin, CMBs are paramagnetic and can be detected with MRI sequences sensitive to magnetic susceptibility (typically, gradient recalled echo T2* weighted images). Nevertheless, their identification remains challenging on T2* magnitude images because of confounding structures and lesions. In this context, T2* phase image may play a key role in better characterizing CMBs because of its direct relationship with local magnetic field variations due to magnetic susceptibility difference. To address this issue, susceptibility-based imaging techniques were proposed, such as Susceptibility Weighted Imaging (SWI) and Quantitative Susceptibility Mapping (QSM). But these techniques have not yet been validated for 2D clinical data in multicenter settings. Here, we introduce 2DHF, a fast 2D phase processing technique embedding both unwrapping and harmonic filtering designed for data acquired in 2D, even with slice-to-slice inconsistencies. This method results in internal field maps which reveal local field details due to magnetic inhomogeneity within the region of interest only. This technique is based on the physical properties of the induced magnetic field and should yield consistent results. A synthetic phantom was created for numerical simulations. It simulates paramagnetic and diamagnetic lesions within a "brain-like' tissue, within a background. The method was evaluated on both this synthetic phantom and multicenter 2D datasets acquired in a standardized clinical setting, and compared with two state-of-the-art methods. It proved to yield consistent results on synthetic images and to be applicable and robust on patient data. As a proof-of-concept, we finally illustrate that it is possible to find a magnetic signature of CMBs and CMCs on internal field maps generated with 2DHF on 2D clinical datasets that gives consistent results with CT-scans in a subsample of 10 subjects acquired with both modalities. See Fig. 6

More details in [16].

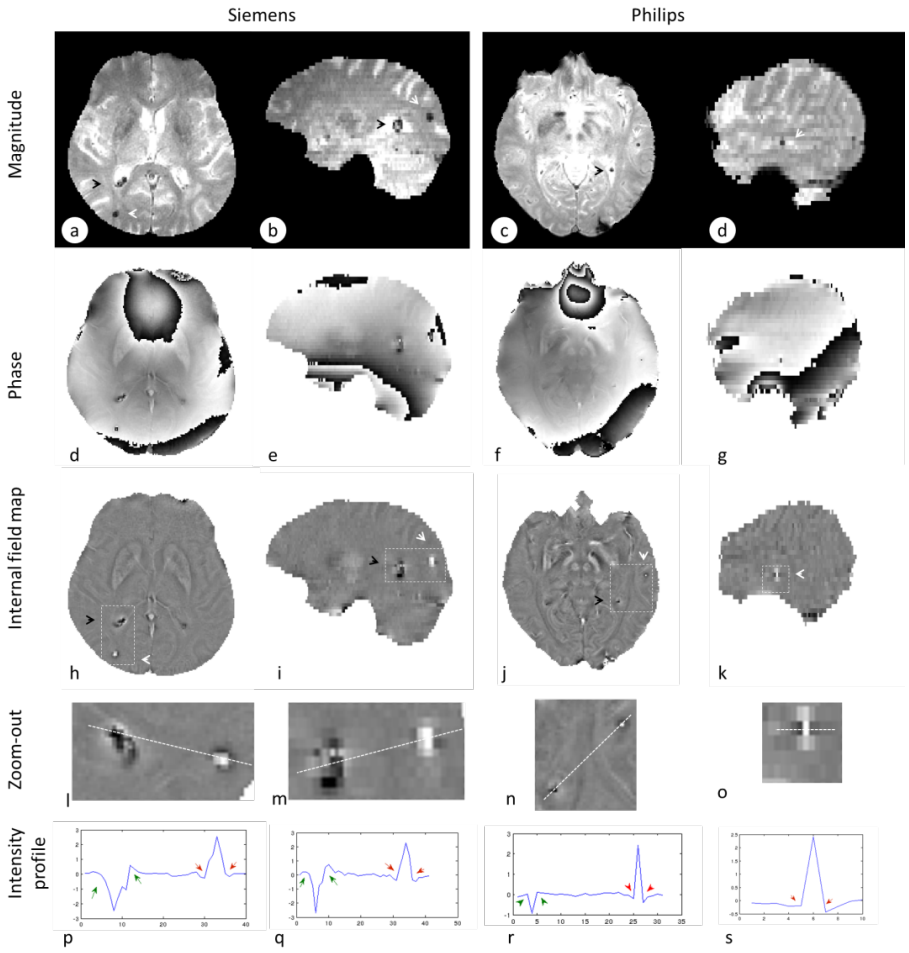


Figure 6. Siemens (left) and Philips (right) axial and sagittal views. Magnitude image (first row), native phase image (second row) and internal field map (third row). Fourth row shows a zoomed out region corresponding to the white rectangle showing CMB with a dipolar pattern (white arrow) and a physiologic calcification of the choroid plexus (black arrow). Note that panel l was rotated. A 1D intensity profile calculated through CMBs and calcification in the zoomed region is displayed in the last row. Note the intensity sign inversion for both side of CMBs (red arrow head), and the calcification (green arrow head). Double headed arrows on panels (l-o) indicate the location of the lines used to generate the intensity profiles.

CLIME Project-Team

6. New Results

6.1. Highlights of the Year

BEST PAPER AWARD :
[20] **VISAPP - International Conference on Computer Vision Theory and Applications**. D. BÉRÉZIAT, I. HERLIN.

6.2. State estimation: analysis and forecast

One major objective of Clime is the conception of new methods of data assimilation in geophysical sciences. Clime is active on several challenging aspects: non-Gaussian assumptions, multiscale assimilation, minimax filtering, etc.

6.2.1. An iterative ensemble Kalman smoother

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

The iterative ensemble Kalman filter (IEnKF) was proposed for improving the performance of the ensemble Kalman filter on 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). IEnKS is an ensemble variational method. It does not require the use of the tangent of the evolution and observation models, nor the adjoint of these models: the required sensitivities (gradient and Hessian) are computed from the ensemble. Looking for the optimal performance, we consider a quasi-static algorithm, out of the many possible extensions. IEnKS was explored on the Lorenz'95 model and on a 2D turbulence model. As a logical extension of IEnKF, IEnKS significantly outperforms the 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.

6.2.2. Modeling and assimilation of lidar signals

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

In this study, we investigate the ability of the chemistry transport model (CTM) Polair3D of the air quality platform Polyphemus to simulate lidar backscattered profiles from model aerosol concentration outputs. This investigation is an important pre-processing stage of data assimilation (validation of the observation operator). To do so, simulated lidar signals are compared to hourly lidar observations performed during the MEGAPOLI (Megacities: Emissions, urban, regional and Global Atmospheric POLlution and climate effects, and Integrated tools for assessment and mitigation) summer experiment in July 2009, when a ground-based mobile lidar was deployed around Paris on-board a van. The comparison is performed for six days (1, 4, 16, 21, 26 and 29 July 2009), corresponding to different levels of pollution and different atmospheric conditions. Overall, Polyphemus reproduces well the vertical distribution of lidar signals and their temporal variability, especially for 1, 16, 26 and 29 July 2009. Discrepancies on 4 and 21 July 2009 are due to high-altitude aerosol layers, which are not well modeled. In the second part of this study, two new algorithms for assimilating lidar observations based on the optimal interpolation method are presented. One algorithm

analyses PM_{10} (particulate matter with diameter less than 10 μm) concentrations. Another analyses $PM_{2.5}$ (particulate matter with diameter less than 2.5 μm) and $PM_{2.5-10}$ (particulate matter with a diameter higher than 2.5 μm and lower than 10 μm) concentrations separately. The aerosol simulations without and with lidar Data Assimilation (DA) are evaluated using the Airparif (a regional operational network in charge of air quality survey around the Paris area) database to demonstrate the feasibility and usefulness of assimilating lidar profiles for aerosol forecasts. The evaluation shows that lidar DA is more efficient at correcting PM_{10} than $PM_{2.5}$, probably because $PM_{2.5}$ is better modeled than PM_{10}. Furthermore, the algorithm which analyzes both $PM_{2.5}$ and $PM_{2.5-10}$ provides the best scores for PM_{10}. The averaged root-mean-square error (RMSE) of PM_{10} is 11.63 $\mu g\, m^{-3}$ with DA ($PM_{2.5}$ and $PM_{2.5-10}$), compared to 13.69 $\mu g\, m^{-3}$ with DA (PM_{10}) and 17.74 $\mu g\, m^{-3}$ without DA on 1 July 2009. The averaged RMSE of PM_{10} is 4.73 $\mu g\, m^{-3}$ with DA ($PM_{2.5}$ and $PM_{2.5-10}$), against 6.08 $\mu g\, m^{-3}$ with DA (PM_{10}) and 6.67 $\mu g\, m^{-3}$ without DA on 26 July 2009.

6.2.3. Assimilation of lidar signals: application to aerosol forecasting

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

This study represents a new application of assimilating lidar signals to aerosol forecasting. It aims at investigating the impact of a ground-based lidar network on the analysis and short-term forecasts of aerosols through a case study in the Mediterranean basin. To do so, we employ a Data Assimilation (DA) algorithm based on the optimal interpolation method developed in the Polair3D chemistry transport model (CTM) of the Polyphemus air quality modeling platform. We assimilate hourly averaged normalized range-corrected lidar signals retrieved from a 72 h period of intensive and continuous measurements performed in July 2012 by ground-based lidar systems of the European Aerosol Research Lidar Network (EARLINET). Particles with an aerodynamic diameter lower than 2.5 μm ($PM_{2.5}$) and those with an aerodynamic diameter higher than 2.5 μm but lower than 10 ($PM_{10-2.5}$) are analyzed separately using the lidar observations at each DA step. First, we study the spatial and temporal influences of the assimilation of lidar signals on aerosol forecasting. We conduct sensitivity studies on algorithmic parameters, e.g. the horizontal correlation length (L_{h}) used in the background error covariance matrix (50 km, 100 km or 200 km), the altitudes at which DA is performed (0.75–3.5 km, 1.0–3.5 km or 1.5–3.5 km) and the assimilation period length (12 h or 24 h). We find that DA with $L_{\mathrm{h}} = 100$ km and assimilation from 1.0 to 3.5 km during a 12 h assimilation period length leads to the best scores for PM_{10} and $PM_{2.5}$ during the forecast period with reference to available measurements from surface networks. Secondly, the aerosol simulation results without and with lidar DA using the optimal parameters (L_{h} = 100 km, an assimilation altitude range from 1.0 to 3.5 km and a 12 h DA period) are evaluated using the level 2.0 (cloud-screened and quality-assured) aerosol optical depth data from AERONET, and mass concentration measurements (PM_{10} or $PM_{2.5}$) from the French air quality (BDQA) network and the EMEP-Spain/Portugal network. The results show that the simulation with DA leads to better scores than the one without DA for $PM_{2.5}$, PM_{10} and aerosol optical depth. Additionally, the comparison of model results to evaluation data indicates that the temporal impact of assimilating lidar signals is longer than 36 h after the assimilation period.

Fig. 2 shows the performance of assimilating real lidar data over the Mediterranean sea with a view to forecast particulate matter over France.

6.2.4. Local ensemble transform Kalman filter for adaptive optics on extremely large telescopes

Participants: Morgan Gray [LAM, France], Cyril Petit [ONERA, France], Sergei Rodionov [LAM, France], Marc Bocquet, Laurent Bertino [NERSC, Norway], Marc Ferrari [LAM, France], Thierry Fusco [LAM and ONERA, France].

We proposed a new algorithm for an adaptive optics system control law, based on the Linear Quadratic Gaussian approach and a Kalman Filter adaptation with localizations. It allows to handle non-stationary behaviors, to obtain performance close to the optimality defined with the residual phase variance minimization criterion, and to reduce the computational burden with an intrinsically parallel implementation on the Extremely Large Telescopes.

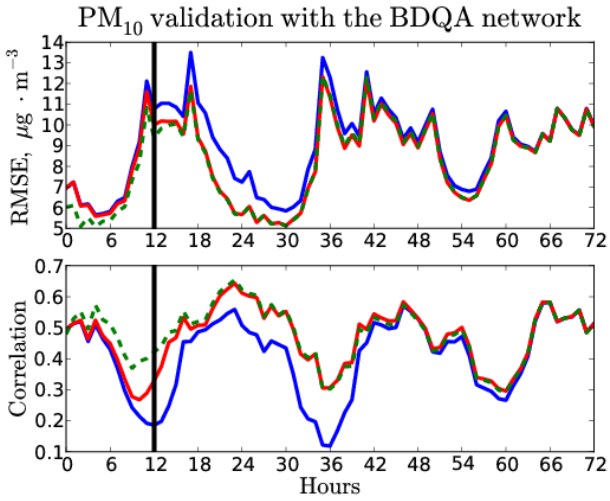


Figure 2. Validation of forecasts of particulate matter PM_{10} using ground stations over France when lidar data have been assimilated over the Mediterranean sea. These forecasts (red line: 12-hour assimilation period and dashed green line: 24-hour assimilation period) are compared to a free run (blue line).

6.3. Inverse modeling

Research on inverse modeling techniques is a major component of Clime, with a focus, in 2014, on hyperparameter estimation when the statistics are non-Gaussian.

6.3.1. Estimation of the caesium-137 source term from the Fukushima Daiichi plant

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 emphasized 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. 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.

6.4. Image assimilation

Sequences of images, such as satellite acquisitions, display structures evolving in time. This information is recognized of major interest by forecasters (meteorologists, oceanographers, etc.) in order to improve the information provided by numerical models. However, the satellite images are mostly assimilated in geophysical models on a point-wise basis, discarding the space-time coherence visualized by the evolution of structures such as clouds. Assimilating in an optimal way image data is of major interest and this issue should be considered in two ways:

- from the model's viewpoint, the location of structures on the observations is used to control the state vector.
- from the image's viewpoint, a model of the dynamics and structures is built from the observations.

6.4.1. 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 depends simultaneously on the initial motion field, at the beginning of the studied temporal window, and on the error value at each time step. The result allows to assess the model error and analyze its impact on motion estimation. The approach is used to estimate geophysical forces (gravity, Coriolis, diffusion) from images in order to better assess the surface dynamics [20] and forecast the displacement of structures like oilspill.

6.4.2. 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 .

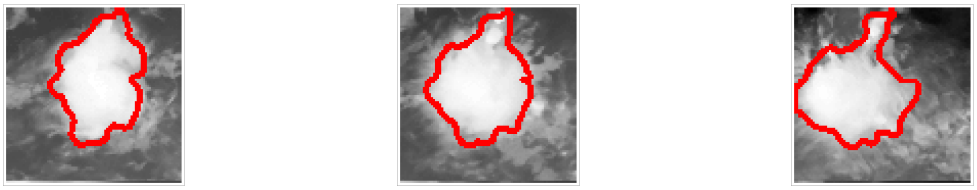


Figure 3. Tracking a tropical cloud. Frames 3, 9, 18 of the sequence.

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 .

Data assimilation is performed either with a 4D-Var variational approach or with a Kalman ensemble method [22]. In the last case, the initial ensemble is obtained from a set of optical flow methods of the literature with various parameters values.

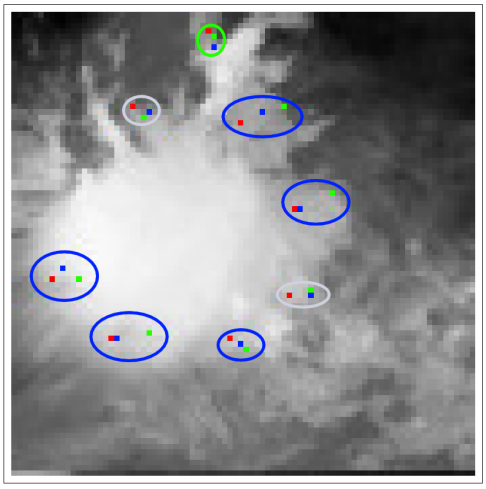


Figure 4. Red point: ground truth. Blue point: our method. Green point: Sun's optical flow. Blue ellipse: our method is the best. Green ellipse: Sun's result is the best. Grey ellipse : results are equivalent.

6.4.3. Motion estimation from images with a waveforms reduced model

Participants: Etienne Huot, Isabelle Herlin, Giuseppe Papari [CFLIR, Belgium].

Dimension reduction is applied to a model of image evolution, composed of transport of velocity and image brightness. Waveform bases are obtained on the image domain for subspaces of images and motion fields, as eigenvectors of previously defined quadratic functions. Image assimilation with the reduced model allows to estimate velocity fields satisfying the space-time properties chosen defined by the user for designing the quadratic function. This approach allows complex geographical domains and suppresses the difficulty of boundary conditions on such domains: these boundary conditions are automatically applied on the bases elements. Motion estimation is then obtained with a reduced model whose state vector is composed of a few components for motion and images. This has to be compared with the initial motion estimation problem that involves a state vector that has a size proportional to the image domain. Current research concern the definition of new quadratic functions from image properties.

6.4.4. Applying POD on a model output dabase for defining a reduced motion model

Participants: Etienne Huot, Isabelle Herlin.

Dimension reduction may also be studied by determining a small size reduced basis obtained by Proper Orthogonal Decomposition (POD) of a motion fields database. This database is constructed for characterizing accurately the surface circulation of the studied area, so that linear combinations of the basis elements obtained by POD accurately describe the motion function observed on satellite image sequences. The database includes the geostrophic motion fields obtained from Sea Level Anomaly reanalysis maps that are available from the MyOcean European project website (http://www.myocean.eu/). Fig. 5 displays such SLA maps and the associated motion fields.

Image assimilation with the POD reduced model allows estimating motion as displayed on Fig. 6 .

6.4.5. Rain nowcasting from radar image acquisitions

Participants: Yann Lepoittevin, Isabelle Herlin.

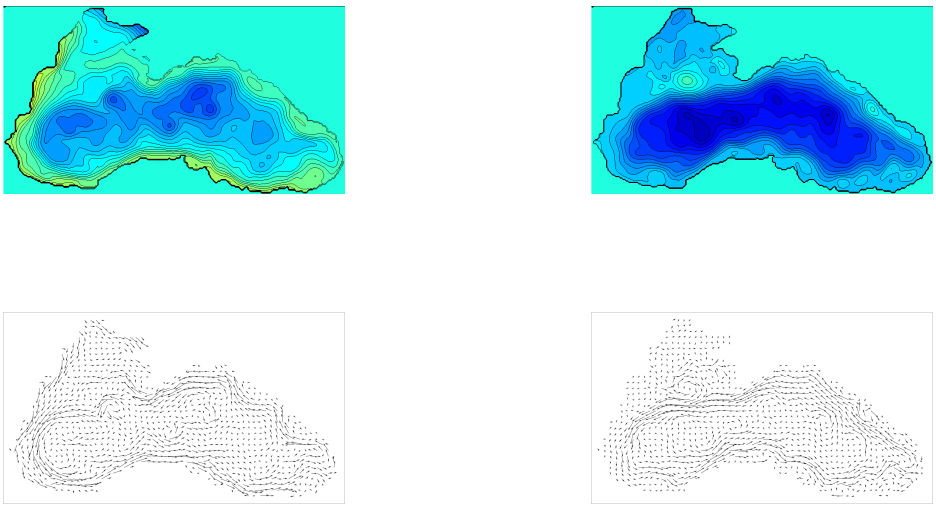


Figure 5. Top: reanalysis of SLA. Bottom: geostrophic motion.

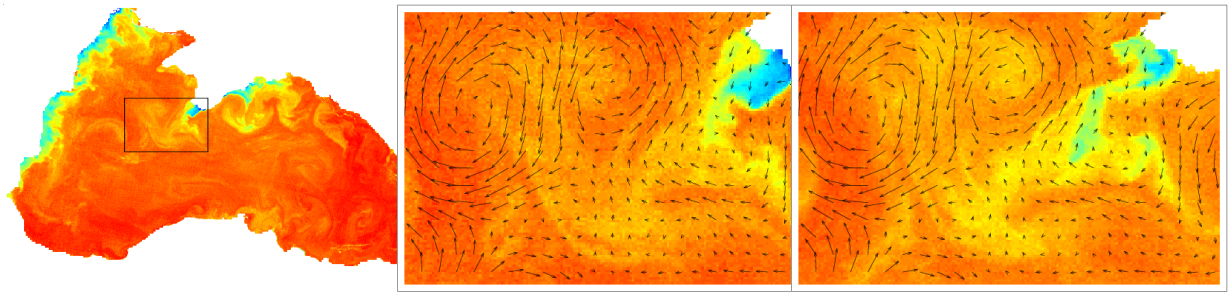


Figure 6. Zoom on a region of interest and motion estimation superposed on two consecutive images.

This research concerns the design of an operational method for rainfall nowcasting that aims at prevention of flash floods. The nowcasting method is based on two main components:

- a data assimilation method, based on radar images, estimates the state of the atmosphere: this is the estimation phase.
- a forecast method uses this estimation to extrapolate the state of the atmosphere in the future: this is the forecast phase.

Results were analyzed by Numtech (partner of a joint I-lab) on space-time neighborhood in order to prevent consequences of flash floods on previously defined zone.

Current research concerns the use of object components in the state vector in order to get an improved motion estimation and a better localization of endangered regions.

6.5. Uncertainty quantification and risk assessment

The uncertainty quantification of environmental models raises a number of problems due to:

- the dimension of the inputs, which can easily be 10^5-10^8 at every time step;
- the dimension of the state vector, which is usually 10^5-10^7;
- the high computational cost required when integrating the model in time.

While uncertainty quantification is a very active field in general, its implementation and development for geosciences requires specific approaches that are investigated by Clime. The project-team tries to determine the best strategies for the generation of ensembles of simulations. In particular, this requires addressing the generation of large multimodel ensembles and the issue of dimension reduction and cost reduction. The dimension reduction consists in projecting the inputs and the state vector to low-dimensional subspaces. The cost reduction is carried out by emulation, i.e., the replacement of costly components with fast surrogates.

6.5.1. Application of sequential aggregation to meteorology

Participants: Jean Thorey, Paul Baudin, Vivien Mallet, Stéphanie Dubost [EDF R&D], Christophe Chaussin [EDF R&D], Laurent Dubus [EDF R&D], Luc Musson-Genon [CEREA, EDF R&D], Laurent Descamps [Météo France], Philippe Blanc [Armines], Gilles Stoltz [CNRS].

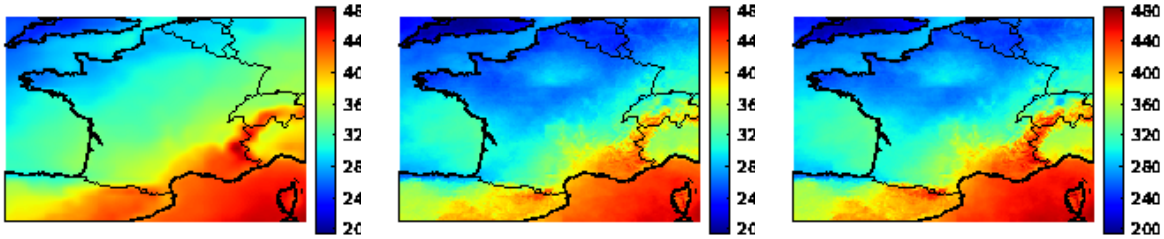
Nowadays, it is standard procedure to generate an ensemble of simulations for a meteorological forecast. Usually, meteorological centers produce a single forecast, out of the ensemble forecasts, computing the ensemble mean (where every model receives an equal weight). It is however possible to apply aggregation methods. When new observations are available, the meteorological centers also compute analyses. Therefore, we can apply the ensemble forecast of analyses. Ensembles of forecasts for mean sea level pressure, from the THORPEX Interactive Grand Global Ensemble, were aggregated with a forecast error decrease by 20% compared to the ensemble mean.

We studied the aggregation of ensembles of solar radiations in the context of photovoltaic production. The observations are based on MeteoSat Second Generation (MSG) and provided by the HelioClim-3 database as gridded fields. The ensembles of forecasts are from the THORPEX Interactive Grand Global Ensemble. The aggregated forecasts show a 20% error decrease compared to the individual forecasts. They are also able to retrieve finer spatial patterns than the ones found in the individual forecasts (see Figure 7).

6.5.2. Sequential aggregation with uncertainty estimation

Participants: Vivien Mallet, Jean Thorey, Paul Baudin, Gilles Stoltz [CNRS].

An important issue is the estimation of the uncertainties associated with the aggregated forecasts. We devised a new approach to predict a probability density function or cumulative distribution function instead of a single aggregated forecast. In practice, the aggregation procedure aims at forecasting the cumulative distribution function of the observations which is simply a Heaviside function centered at the observed value. Our forecast is the weighted empirical cumulative distribution function based on the ensemble of forecasts. The method guarantees that, in the long run, the forecast cumulative distribution function has a continuous ranked probability score at least as good as the best weighted empirical cumulative function with weights constant in time.



(a) (b) (c)

Figure 7. Yearly average of the map of downward shortwave solar radiation in Wm^{-2}*, for an ensemble mean (a), for our aggregated forecasts (b) and observed (c).*

6.5.3. Sensitivity analysis in the dispersion of radionuclides

Participants: Sylvain Girard, 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. The sensitivity analysis was carried out with the Morris method and by computing Sobol' indices. Both approaches were found to be consistent. Computing the Sobol' indices required the use of Gaussian process emulation, which proved to be successful at least on targets averaged in time and space.

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 in time and space (see Figure 8). 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.

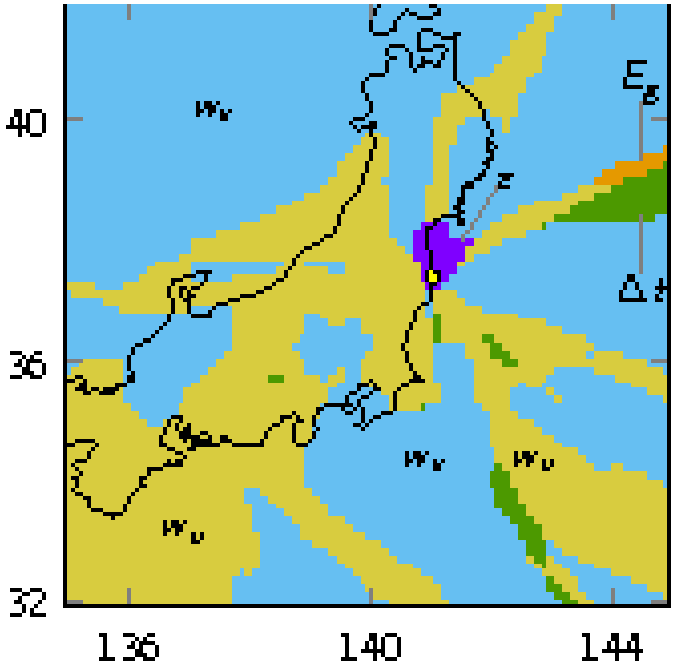


Figure 8. Variables that influence the most the atmospheric radioactivity after Fukushima disaster. z is the emissions altitude; Δt is the time shift on emissions; E_g stands for the emissions of noble gas; w_u and w_v are for zonal and meridional winds, respectively.

LIFEWARE Team

6. New Results

6.1. Highlight: Xavier Duportet laureate of the AEF docteurs-entrepreneurs prize

Xavier Duportet has been awarded the AEF prize at the docteurs-entrepreneurs competition. His thesis, made jointly within Lifeware and the Weiss lab at MIT, was entitled "Developing new tools and platforms for mammalian synthetic biology: from the assembly and chromosomal integration of large genetic circuits to the engineering of artificial intercellular communication systems". He published his research in *Nucleic Acids Research* and *Nature Biotechnology* [7], [5]. In particular, he demonstrated the assembly and chromosomal integration in mammalian cells of the largest gene circuit integrated to date. Subsequently, he co-founded the startup company PhageX. He was also a laureate of the Concours National de Création d'Entreprises Innovantes and the Concours Mondial d'Innovation (personalized medicine track).

He is the president of the Hello Tomorrow challenge and vice-president of the Osons La France initiative. He has notably been featured in articles published in Le Monde, L'Obs, and L'Opinion. He has been an invited speaker at the prestigious 4th Congreso del Futuro in Santiago (Chili). [7]

6.2. Highlight: François Fages laureate of the French Academy of Sciences

François Fages was very honoured to receive the Michel Monpetit prize 2014 of the French Academy of Sciences for his contributions to fundamental computer science (unification theory and constraint logic programming) and computational systems biology (modeling of biochemical networks and design and supervision of the implementation of the BIOCHAM software).

6.3. Highlight: Pauline Traynard Best Student Paper Prize at CMSB 2014, for Trace Simplifications preserving Temporal Logic Formulae with Case Study in a Coupled Model of the Cell Cycle and the Circadian Clock

Participants: François Fages, Sylvain Soliman, Pauline Traynard.

Pauline Traynard was very pleased to receive the Best Student Paper Prize of the twelth International Conference on Computational Methods for Systems Biology, 17-19 November 2014, Univ. of Manchester, UK, for a communication on trace simplifications preserving temporal logic properties [19].

Calibrating dynamical models on experimental data time series is a central task in computational systems biology. When numerical values for model parameters can be found to fit the data, the model can be used to make predictions, whereas the absence of any good fit may suggest to revisit the structure of the model and gain new insights in the biology of the system. Temporal logic provides a formal framework to deal with imprecise data and specify a wide variety of dynamical behaviors. It can be used to extract information from numerical traces coming from either experimental data or model simulations, and to specify the expected behaviors for model calibration. The computation time of the different methods depends on the number of points in the trace so the question of trace simplification is important to improve their performance. In [19] we study this problem and provide a series of trace simplifications which are correct to perform for some common temporal logic formulae. We give some general soundness theorems, and apply this approach to period and phase constraints on the circadian clock and the cell cycle. In this application, temporal logic patterns are used to compute the relevant characteristics of the experimental traces, and to measure the adequacy of the model to its specification on simulation traces. Speed-ups by several orders of magnitude are obtained by trace simplification even when produced by smart numerical integration methods.

6.4. Highlight: Modeling Dynamics of Cell-to-Cell Variability in TRAIL-induced Apoptosis Explains Fractional Killing and Predicts Reversible Resistance

Participants: Grégory Batt, François Bertaux, Szymon Stoma.

Isogenic cells sensing identical external signals can take markedly different decisions. Such decisions often correlate with pre-existing cell-to-cell differences in protein levels. When not neglected in signal transduction models, these differences are accounted for in a static manner, by assuming randomly distributed initial protein levels. However, this approach ignores the *a priori* non-trivial interplay between signal transduction and the source of this cell-to-cell variability: temporal fluctuations of protein levels in individual cells, driven by noisy synthesis and degradation. Thus, modeling protein fluctuations, rather than their consequences on the initial population heterogeneity, would set the quantitative analysis of signal transduction on firmer grounds. Adopting this dynamical view on cell-to-cell differences amounts to recast extrinsic variability into intrinsic noise. In collaboration with Dirk Drasdo (EPI Mmaba), we proposed a generic approach to merge, in a systematic and principled manner, signal transduction models with stochastic protein turnover models. When applied to an established kinetic model of TRAIL-induced apoptosis, our approach markedly increased model prediction capabilities [4]. We obtained a mechanistic explanation of yet-unexplained observations on fractional killing and non-trivial robust predictions of the temporal evolution of cell resistance to TRAIL in HeLa cells. Our results provide an alternative explanation to survival via induction of survival pathways since no TRAIL-induced regulations are needed and suggest that short-lived anti-apoptotic protein Mcl1 exhibit large and rare fluctuations. More generally, our results highlight the importance of accounting for stochastic protein turnover to quantitatively understand signal transduction over extended durations, and imply that fluctuations of short-lived proteins deserve particular attention. [4]

6.5. Towards Real-time Control of Gene Expression at the Single Cell Level: A Stochastic Control Approach

Participants: Grégory Batt, Pascal Hersen.

Recent works have demonstrated the experimental feasibility of real-time gene expression control based on deterministic controllers. By taking control of the level of intracellular proteins, one can probe single-cell dynamics with unprecedented flexibility. However, single-cell dynamics are stochastic in nature, and a control framework explicitly accounting for this variability is presently lacking. In [21], we devised a stochastic control framework, based on Model Predictive Control, which fills this gap.

Based on a stochastic modelling of the gene response dynamics, our approach combined a full state-feedback receding-horizon controller with a real-time estimation method that compensated for unobserved state variables. Using previously developed models of osmostress-inducible gene expression in yeast, we showed *in silico* that our stochastic control approach outperformed deterministic control design in the regulation of single cells. This contribution lead to envision the application of the proposed framework to wet lab experiments in yeast.

This work was done in collaboration with Alfonso Carta (EPI BIOCORE), Eugenio Cinquemani (EPI IBIS), Lakshmeesh Maruthi and Ilya Tkachev (TU Delft), and Alessandro Abate (Oxford U).

6.6. A Platform for Rapid Prototyping of Synthetic Gene Networks in Mammalian Cells

Participants: Grégory Batt, Xavier Duportet, Pascal Hersen.

Mammalian synthetic biology may provide novel therapeutic strategies, help decipher new paths for drug discovery and facilitate synthesis of valuable molecules. Yet, our capacity to genetically program cells is currently hampered by the lack of efficient approaches to streamline the design, construction and screening of synthetic gene networks. To address this problem, we developed a framework for modular and combinatorial assembly of functional (multi)gene expression vectors and showed their efficient and specific targeted integration into a well-defined chromosomal context in mammalian cells.

In [7], in collaboration with the Weiss lab and the MSC lab, we demonstrated the potential of this framework by assembling and integrating different functional mammalian regulatory networks including the largest gene circuit built and chromosomally integrated to date (6 transcription units, 27kb), encoding an inducible memory device. Using a library of 18 different circuits as a proof of concept, we also demonstrated that our method enabled one-pot/single-flask chromosomal integration and screening of circuit libraries. This rapid and powerful prototyping platform is well suited for comparative studies of genetic regulatory elements, genes and multi-gene circuits as well as facile development of libraries of isogenic engineered cell lines.

6.7. Reconfigurable Circuitry in Biochemical Systems

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

Realizing complex systems within a biochemical environment is a common pursuit in synthetic biology. Such systems achieve certain computation through properly designed biochemical reactions. Despite fruitful progress being made, most existing reaction designs have fixed target functionality. Their lack of reconfigurability can be disadvantageous, especially when a system has to adapt to a varying biochemical environment.

When control systems are of concern, linear control is one of the most widely applied control methods. Any linear control system can be realized with three elementary building blocks: integration, gain, and summation. Realizing linear control with biochemical reactions has been proposed in previous work, where reaction rates of the underlying reactions play a key role to achieve the desired building blocks. Essentially the reaction rates have to be matched exactly, and it imposes serious practicality restriction because in reality the reaction rates of available reactions are predetermined and can be limited. In [16] we devise a mechanism to make linear control systems configurable by adding auxiliary species as control knobs. The concentrations of the auxiliary species can be adjusted not only to compensate reaction rate mismatch, but also to reconfigure different control systems out of the same control architecture.

Furthermore, in [15] we propose an analog approach to economically construct a reconfigurable logic circuit similar to a silicon based field programmable gate array (FPGA). The effective “logic” and “interconnect” of the circuit can be dynamically reconfigured by controlling the concentrations of certain knob species. We study a potential biomedical application of our reconfigurable circuitry to disease diagnosis and therapy at a molecular level.

6.8. Inferring Reaction Systems from Ordinary Differential Equations

Participants: François Fages, Steven Gay, Sylvain Soliman.

In Mathematical Biology, many dynamical models of biochemical reaction systems are presented with Ordinary Differential Equations (ODE). Once kinetic parameter values are fixed, this simple mathematical formalism completely defines the dynamical behavior of a system of biochemical reactions and provides powerful tools for deterministic simulations, parameter sensitivity analysis, bifurcation analysis, etc. However, without requiring any information on the reaction kinetics and parameter values, various qualitative analyses can be performed using the structure of the reactions, provided the reactants, products and modifiers of each reaction are precisely defined. In order to apply these structural methods to parametric ODE models, we study a mathematical condition for expressing the consistency between the structure and the kinetics of a reaction, without restricting to Mass Action law kinetics. This condition, satisfied in particular by standard kinetic laws, entails a remarkable property of independence of the influence graph from the kinetics of the reactions. We derive from this study a heuristic algorithm which, given a system of ODEs as input, computes a system of reactions with the same ODE semantics, by inferring well-formed reactions whenever possible. We show how

this strategy is capable of automatically curating the writing of ODE models in SBML, and present some statistics obtained on the model repository biomodels.net [8].

6.9. Model Reductions by Tropical Equilibration

Participants: François Fages, Sylvain Soliman.

Model reduction is a central topic in systems biology and dynamical systems theory, for reducing the complexity of detailed models, finding important parameters, and developing multi-scale models for instance. While singular 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 methods is in the computation of tropical equilibrations. In [11] 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 detailed reduction of a simple biochemical mechanism, the Michaelis-Menten enzymatic reaction model, and second, with large-scale performance figures obtained on the http://biomodels.net webcite repository.

6.10. Model Reductions by Subgraph Epimorphisms

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

In [9] we follow another route based on a purely structural method and study the 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 vertex deletion and merge operations that transforms a first reaction graph into a second graph. 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. In this paper, we study theoretical properties of subgraph epimorphisms in general directed graphs. We first characterize subgraph epimorphisms (SEPI), subgraph isomorphisms (SISO) and graph epimorphisms (EPI) in terms of graph transformation operations. Then we study the graph distance measures induced by these transformations. We show that they define metrics on graphs and compare them. On the algorithmic side, we show that the SEPI existence problem is NP-complete by reduction of SAT, and present a constraint satisfaction algorithm that has been successfully used to solve practical SEPI problems on a large benchmark of reaction graphs from systems biology.

6.11. Temporal Logic Modeling of Dynamical Behaviors: First-Order Patterns and Solvers

Participants: François Fages, Pauline Traynard, Sylvain Soliman.

We have written a book chapter [20] to describe how quantitative temporal logic formulae can be used to formalize imprecise dynamical behaviors of biological systems, and how such a formal specification of experimental observations can be used to calibrate models to real data, in a more versatile way than with curve fitting algorithms, and with more efficient dedicated solvers than with generic temporal logic solvers.

Based on this article, we investigated the correctness of various trace simplification methods, as mentionned in the highlight section above [19].

6.12. Logical Modeling of the Mammalian Cell Cycle

Participants: François Fages, Pauline Traynard, Denis Thieffry.

The molecular networks controlling cell cycle progression in various organisms have been previously modelled, predominantly using differential equations. However, this approach meets various difficulties as one tries to include additional regulatory components and mechanisms. This led to the development of qualitative dynamical models based on Boolean or multilevel frameworks, which are easier to define, simulate, analyse and compose. In a poster presented at ECCB 2014, we revisit the Boolean model of Fauré et al. for the core network controlling G/S transition in mammalian cell cycle, taking into account recent advances in the characterisation of the underlying molecular networks to obtain a better qualitative consistency between model simulations and documented mutants features. In particular, we introduced Skp2, the substrate recruiting component of the SCFSkp2 complex, which targets cell cycle control elements, such as p27, and is repressed by the tumour suppressor protein Rb. Furthermore, to supersede the limitations inherent to the Boolean simplifications, we have considered the association of multilevel logical components with key cell cycle regulators, including the tumour suppressor protein Rb. Indeed, it is well established that differently phosphorylated forms of Rb result in different effects on other components of the network, which can be faithfully modelled using a multilevel rather than a Boolean variable. To evaluate the dynamical properties of the resulting models, we perform synchronous and asynchronous simulations using the software GINsim (http://www.ginsim.org), for both the wild-type case and documented perturbations (e.g. combinations of loss- or gain-of-function mutations). In addition, we have designed a series of temporal logic queries (expressed in the CTL language), which enable an efficient and automatic verification of key dynamical properties (existence of a cyclic attractor or of a stable state, conditions on the order of changes of component levels, etc.), using the popular symbolic model checker NuSMV. This strategy greatly facilitates the dynamical analysis of increasingly detailed and complex cell cycle models. Our goal is to obtain a core cell cycle model consistent with the most relevant experimental results on mammalian cells, which will then be used as a module in more comprehensive cellular models, including cross-talks with the circadian clock network and key signalling pathways, whose deregulation underlies the development of various cancers.

6.13. A Greedy Heuristic for Optimizing Metro Regenerative Energy Usage compared to CMA-ES and MILP

Participants: François Fages, David Fournier.

When the regenerative braking energy cannot be stored by the metro producing it, it has to be used instantaneously on the network, otherwise it is lost. In this case, the accelerating and braking trains need be synchronized to fully benefit from the regenerative energy, and a metro timetable is energetically optimized when all the regenerative braking is utilized to power other trains. This synchronization consists in lining up each braking train with an accelerating one in its neighbourhood. Doing so, the latter will benefit from the regenerative energy of the former. In [17], [3] a fast greedy heuristic is proposed to tackle the problem of minimizing the energy consumption of a metro timetable by modifying solely the dwell times in stations. This heuristic is compared to a state-of-the-art meta heuristic called the covariance matrix adaptation evolution strategy (CMA-ES) and shows similar results with much faster computation time. Finally, it is shown that a run of the algorithm on a full timetable may reduce its energy consumption by 5.1%.

MAMBA Team

6. New Results

6.1. Highlights of the Year

Benoît Perthame was invited as plenary speaker for the International Congress of Mathematicians ICM 2014 (Seoul, http://www.icm2014.org), that attracted more than 5000 participants. This is the first time that a mathematician working in mathematics applied to biology was invited at ICM, which is the most prestigious conference for mathematicians of all fields. This represents a consecration both for Benoît Perthame's work and for the MAMBA team, and more generally for the whole domain of mathematics applied to biology.

Marie Doumic was a plenary speaker at the ECMTB 2014 (Göteborg, http://ecmtb2014.org/ 600 participants).

Dirk Drasdo was invited speaker at the Systems Biology of Human Diseases conference (Harvard University, http://www.csb2.org/events/sbhd-2014).

Five articles are noteworthy in terms of bibliometry:

• (*Impact factor 11.2*) F. SCHLIESS, S. HOEHME, S. HENKEL, A. GHALLAB, D. DRIESCH, J. BÖTTGER, R. GUTHKE, M. PFAFF, J. HENGSTLER, R. GEBHARDT, D. HÄUSSINGER, D. DRASDO, S. ZELLMER. Integrated metabolic spatial-temporal model for the prediction of ammonia detoxification during liver damage and regeneration, *Hepatology*, Dec. 2014, vol. 60, no 6, pp. 2040-2051, https://hal.inria.fr/hal-01110646 [17]

• (*Impact factor 10.4*) D. DRASDO, S. HOEHME, J. G. HENGSTLER. How predictive quantitative modeling of tissue organization can inform liver disease pathogenesis, *Journal of Hepatology*, Oct. 2014, vol. 61, no 4, pp. 951-956 [DOI : 10.1016/J.JHEP.2014.06.013], https://hal.inria.fr/hal-01110644 [7]

• (*Impact factor 10.7*) S.R.K. VEDULA, G. PEYRET, I. CHEDDADI, T. CHEN, A. BRUGUÉS, H. HIRATA, H. LOPEZ-MENENDEZ, Y. TOYAMA, L. NEVES DE ALMEIDA, X. TREPAT, C.T. LIM, B. LADOUX. Mechanics of epithelial closure over non-adherent environments, *Nature Communications*, Jan. 2015, vol. 6, art. number 6111[DOI : 10.1038/ncomms7111], http://www.nature.com/ncomms/2015/150122/ncomms7111/abs/ncomms7111.html (open access)

• (*Impact factor 7.5*) L. ROBERT, M. HOFFMANN, N. KRELL, S. AYMERICH, J. ROBERT, M. DOUMIC. Division in Escherichia coli is triggered by a size-sensing rather than a timing mechanism, in "BMC Biology", 2014, vol. 12, no 1, 17 p. [DOI : 10.1186/1741-7007-12-17], https://hal.inria.fr/hal-00981312 [16]

• (*Impact factor 9.3*) R. H. CHISHOLM, T. LORENZI, A. LORZ, A. K. LARSEN, L. ALMEIDA, A. ESCARGUEIL, J. CLAIRAMBAULT. Emergence of drug tolerance in cancer cell populations: an evolutionary outcome of selection, nongenetic instability and stress-induced adaptation, *Cancer Research* (Mathematical oncology), 10p.+suppl. mat., in press, Jan. 2015, https://hal.archives-ouvertes.fr/hal-01111271 [33]

6.2. Cancer

Participants: Luís Lopes Neves de Almeida, José Luis Avila Alonso [DISCO Inria team], Catherine Bonnet [DISCO Inria team], Rebecca Chisholm, Jean Clairambault, François Delhommeau [Hæmatology department, St Antoine Hospital, Paris], Luna Dimitrio [former PhD student and Mamba member], Ján Eliaš, Alexandre Escargueil [Cancer biology and therapeutics lab, St Antoine Hospital, Paris], Pierre Hirsch [Hæmatology department, St Antoine Hospital, Paris], Michal Kowalczyk [Univ. Santiago de Chile], Annette Larsen [Cancer biology and therapeutics lab, St Antoine Hospital, Paris], Tommaso Lorenzi, Alexander Lorz, Anna Marciniak-Czochra [Univ. Heidelberg], Roberto Natalini [IAC-CNR, Univ. Tor Vergata, Rome], Silviu Iulan Niculescu [DISCO Inria team], Hitay Özbay [Bilkent Univ., Ankara], Benoît Perthame, Andrada Maran, Fernando Quirós [Univ. Autónoma de Madrid], Michèle Sabbah [Cancer biology and therapeutics lab, St Antoine Hospital, Paris], Thomas Stiehl [Univ. Heidelberg], Min Tang [Jiaotong University, Shanghai], Emmanuel Trélat [LJLL, UPMC], Nicolas Vauchelet, Romain Yvinec [INRA Tours].

6.2.1. Drug resistance.

We have continued to develop our phenotypically based models of drug-induced drug resistance in cancer cell populations, representing their Darwinian evolution under drug pressure by integro-differential equations. In one of them [40], a 1D space variable has been added to the phenotypic structure variable to account for drug diffusion in tumour spheroids. In another one [33], where deterministic and agent-based modelling are processed in parallel, we have considered a physiologically based 2-dimensional phenotypic structure variable, in order to take account of previously published biological observations on (reversible) drug tolerance persistence in a population of non-small cell lung cancer (NSCLC) cells [0], reproducing the observations and assessing the model by testing biologically based hypotheses. Together with ongoing work with E. Trélat and A. Lorz on drug therapy optimisation, using such phenotype-based models to overcome drug resistance, this has represented a significant part of our work on the subject, which is conducted in close collaboration with the INSERM-UPMC team "Cancer biology and therapeutics" (A. Larsen, A. Escargueil, M. Sabbah) at St Antoine Hospital.

6.2.2. Reversible drug resistance and fractional killing in tumor cell population treatment.

We developed a model of drug resistance in TRAIL (TNF-Related Apoptosis Induced-Ligand) treatment in HeLa cell lines. The TRAIL signal transduction pathway is one of the best studied apoptosis pathways and hence permits detailed comparisons with data. Our model was able to explain experimental observations fractional killing and cell-to-cell variability, and predicted reversible resistance [3]. (Work in close collaboration with G. Batt and S. Stoma from the Inria team LIFEWARE.)

6.2.3. Radiotherapy.

Radiation is still a major treatment in cancer. We explored by extensive computer simulations using an agent-based model the consequences of spatially inhomogeneous irradiation. The model predicted that in the case of different competing sub-populations, namely cancer stem cells with unlimited division capacity, and cancer cells with limited division capacity, inhomogeneous radiation focusing higher doses at the tumour centre and lower doses at the tumour periphery should outperform homogeneous irradiation [12]. Cancer stem cells are believed to have a longer cell cycle duration than cancer cells, and are less radiosensitive than cancer cells, which is why they often survive radiation and lead to tumour relapse.

6.2.4. Intercellular interactions in epithelio-mesenchymal transition (EMT).

A PhD thesis on this subject, co-supervised by L. Almeida and M. Sabbah (INSERM team "Cancer biology and therapeutics", St Antoine) has begun at Fall. It is also based on phenotype-structured modelling of Darwinian evolution in cancer cell populations.

6.2.5. Interactions between tumour cell populations and their cellular micro-environment.

A phenotype-structured model of the interactions between a brest cancer cell population (MCF7 cultured cells, collaboration with M. Sabbah, St Antoine Hospital) and its adipocyte stroma support cell population has been developed (T. Lorenzi, J. Clairambault), which, beyond submitted proposals (ANR, Emergence Paris-Sorbonne Universités call), will be studied and experimentally identified in a forthcoming internship (January-June 2015) and PhD thesis in applied mathematics.

[0] Sharma *et al.*, Cell, April 2010

6.2.6. Hele-Shaw model of tumour growth.

In the growing field of mathematical analysis of mechanical domain of tumor growth, we focus on the rigorous link between cells models, relying on mechanical properties of cells, and free boundary problem, where the tumor is described by the dynamics of its boundary. The latter model is referred to Hele-Shaw model [44]. Benoît Perthame, Min Tang and Nicolas Vauchelet have proved the rigorous derivation of a geometric model of the Hele-Shaw type for a model with viscoelastic forces, constructing analytically traveling wave solutions of the Hele-Shaw model of tumor growth with nutrient that explain theoretically the numerical results observed. The limiting model exhibits travelling waves, which have been investigated in [43]. Another interesting feature for this model is the transversal instability occurring when the spatial dimension is greater than 1. Together with Fernando Quirós (Univ. Autónoma de Madrid), the aforementioned have also formulated a Hele-Shaw type free-boundary problem for a tumor growing under the combined effects of pressure forces, cell multiplication and active motion, the latter being the novelty of this study [61]. In order to understand the emergence of instabilities in the Hele-Shaw model with nutrients, Michal Kowalczyk (Univ. Chile, Santiago), Benoît Perthame and Nicolas Vauchelet have studied a related model of thermo-reactive diffusion where they can study the spectrum of the linearized system around a traveling wave and in which they can compute the transition to instability in terms of a parameter related to the ratio between heat conduction and molecular diffusion. However, the rigorous study of such instabilities for the whole system of equations is not reachable for the moment; only a study for a simplified model has been performed in [39].

6.2.7. Modelling and control of acute myeloblastic leukæmia (AML).

The collaboration with the Disco project-team has been continued, leading to one book chapter [25], four conference proceedings [21], [22], [23], [24] and JL Avila Alonso's PhD thesis defence.

In more detail:
Starting initially from a PDE model of hematopoiesis designed by Adimy *et al.*[0], we have derived several models of healthy or cancer cell dynamics in hematopoiesis and performed several stability analyses.

We have proposed in [25] a new mathematical model of the cell dynamics in acute myeloid leukæmia (AML) which takes into account the four different phases of the proliferating compartment as well the fast self-renewal phenomenon frequently observed in AML. As was the case in [25] this model is transformed into a distributed delay system and was analyzed here with input-output techniques. Local stability conditions for an equilibrium point of interest are derived in terms of a set of inequalities involving the parameters of the mathematical model.

We have also studied a coupled delay model for healthy and cancer cell dynamics in AML consisting of two stages of maturation for cancer cells and three stages of maturation for healthy cells. For a particular healthy equilibrium point, locally stability conditions involving the parameters of the mathematical model have been obtained [22], [23].

We have performed in [21] a stability analysis of both the PDE model of healthy hæmatopoiesis and a coupled PDE model of healthy and cancer cell dynamics. The stability conditions obtained here in the time domain strengthen the idea that fast self-renewal plays an important role in AML.

A time-domain stability analysis by means of Lyapunov-Krasovskii functionals has been performed on the delay system modeling healthy hematopoiesis for a strictly positive equilibrium point of interest.

Furthermore, a working collaboration on AML modelling with Anna Marciniak-Czochra (Univ. Heidelberg) was also initiated by the end of 2014 by a visit of three of us (C. Bonnet, J. Clairambault, T. Lorenzi) to Heidelberg and a visit of T. Stiehl, A. Marciniak-Czochra PhD student, to Paris. The topics we plan to investigate are, beyond the role of fast self renewal in AML cell populations, the part played by clonal heterogeneity in leukæmic cell populations and the issues it raises in therapeutics, a well known clinical problem in clinical hæmatology.

[0] Adimy, M., Crauste, F., El Abllaoui, A. Discrete maturity-structured model of cell differentiation with applications to acute myelogenous leukemia, *J. Biol. Sys.*, 16(3):395-424, 2008

Let us also mention that on the subject of early leukæmogenesis, Andrada Qillas Maran has undertaken a PhD thesis under the supervision of J. Clairambault and B. Perthame. Models relying on piecewise deterministic Markov processes (PDMPs), designed and studied by R. Yvinec (INRA Tours) for the single-cell part of the model under construction, will be used in collaboration with him. Our clinical referents in hæmatology for this PhD work are F. Delhommeau and P. Hirsch (St Antoine Hospital).

6.2.8. *The p53 protein spatio-temporal dynamics.*

The development of our molecular-based model of the spatio-temporal intracellular dynamics of the p53 protein (the so-called "guardian of the genome") has been continued [55], [9], leading us also, more generally, to propose a modelling frame dedicated to the dynamics of intracellular proteins and their gene regulatory networks [8].

6.2.9. *Others.*

In a collaboration with ANGE, B. Perthame has studied a data assimilation algorithm for multidimensional hyperbolic conservation laws using kinetic schemes and kinetic formulations.

6.3. Aggregation kinetics

Participants: Tom Banks, Thibault Bourgeron, Marc Hoffmann, Marie Doumic-Jauffret, Nathalie Krell, Benoît Perthame, Stéphanie Prigent, Human Rezaei, Nathalie Robert, Léon Matar Tine [Univ. Lyon and Dracula Inria team], Jorge Zubelli [IMPA, Rio de Janeiro].

6.3.1. *Time Asymptotics for Fragmentation Equations*

Fragmentation and growth-fragmentation equations is a family of problems with varied and wide applications. This paper is devoted to description of the long time time asymptotics of two critical cases of these equations, when the division rate is constant and the growth rate is linear or zero. The study of these cases may be reduced to the study of the following fragmentation equation:

$$\frac{\partial}{\partial t}u(t,x) + u(t,x) = \int_x^\infty k_0(xy)u(t,y)dy.$$

Using the Mellin transform of the equation, we determine the long time behavior of the solutions. Our results show in particular the strong dependence of this asymptotic behavior with respect to the initial data.

6.3.2. *Estimating the division rate in a size-structured population.*

The problem which was considered in [5] consists in estimating the division rate from the stable size distribution of the population, which is easily measured, but non-smooth. We propose a method based on the Mellin transform for growth-fragmentation equations with self-similar kernels. We build a sequence of functions which converges to the density of the population in division, simultaneously in several weighted L^2 spaces, as the measurement error goes to 0. This improves previous results for self-similar kernels [0] and allows us to understand the partial results for general fragmentation kernels [0]. Numerical simulations confirm the theoretical results. Moreover, our numerical method is tested on real biological data, arising from a bacteria growth and fission experiment.

6.3.3. *What governs bacterial growth? The "sizer" vs the "timer" model*

We applied the previously seen inverse problem methodology [5] to a fundamental biological problem: what governs the bacterial growth?

[0] Perthame and Zubelli, *Inv. Prob.*, 2007
[0] Doumic and Tine, *J. Math. Biol.*, 2012

Many organisms coordinate cell growth and division through size control mechanisms: cells must reach a critical size to trigger a cell cycle event. Bacterial division is often assumed to be controlled in this way, but experimental evidence to support this assumption is still lacking. Theoretical arguments show that size control is required to maintain size homeostasis in the case of exponential growth of individual cells. Nevertheless, if the growth law deviates slightly from exponential for very small cells, homeostasis can be maintained with a simple 'timer' triggering division. Therefore, deciding whether division control in bacteria relies on a 'timer' or 'sizer' mechanism requires quantitative comparisons between models and data.

The timer and sizer hypotheses find a natural expression in models based on partial differential equations. Here we test these models with recent data on single-cell growth of Escherichia coli. We demonstrate that a size-independent timer mechanism for division control, though theoretically possible, is quantitatively incompatible with the data and extremely sensitive to slight variations in the growth law. In contrast, a sizer model is robust and fits the data well. In addition, we tested the effect of variability in individual growth rates and noise in septum positioning and found that size control is robust to this phenotypic noise.

Confrontations between cell cycle models and data usually suffer from a lack of high-quality data and suitable statistical estimation techniques. In the study [16] we had overcome these limitations by using high precision measurements of tens of thousands of single bacterial cells combined with recent statistical inference methods to estimate the division rate within the models. We therefore provided the first precise quantitative assessment of different cell cycle models.

6.3.4. Size distribution of amyloid fibrils. Mathematical models and experimental data.

More than twenty types of proteins can adopt misfolded conformations, which can co-aggregate into amyloid fibrils, and are related to pathologies such as Alzheimer's disease. In [15], we surveyed mathematical models for aggregation chain reactions, and discussed the ability to use them to understand amyloid distributions. Numerous reactions have been proposed to play a role in their aggregation kinetics, though the relative importance of each reaction in vivo is unclear: these include activation steps, with nucleation compared to initiation, disaggregation steps, with depolymerization compared to fragmentation, and additional processes such as filament coalescence or secondary nucleation. We have statistically analysed the shape of the size distribution of prion fibrils, with the specific example of truncated data due to the experimental technique (electron microscopy). A model of polymerization and depolymerization succeeds in explaining this distribution. It is a very plausible scheme though, as evidenced in the review of other mathematical models, other types of reactions could also give rise to the same type of distributions.

To clarify how these fibrils are able to incorporate additional units, prion fibril aggregation and disaggregation kinetics were experimentally studied using Static Light Scattering (SLS) [45]. Values that are functions of $\sum i^2 c_i$ (for $i > 0$) with c_i being the concentration of fibrils of size i, were then measured as a function of time. An initial model, adapted from the Becker-Döring system that considers all fibrils to react similarly is not able to reproduce the observed in vitro behaviour. Our second model involves an additional compartment of fibrils unable to incorporate more prion units. This model leads to kinetic coefficients which are biologically plausible and correctly simulates the first experimental steps for prion aggregation.

In the formation of large clusters out of small particles, the initializing step is called the nucleation, and consists in the spontaneous reaction of agents which aggregate into small and stable polymers called nucleus. After this early step, the polymers are involved into a bunch of reactions such as polymerization, fragmentation and coalescence. Since there may be several orders of magnitude between the size of a particle and the size of an aggregate, building efficient numerical schemes to capture accurately the kinetics of the reaction is a delicate step of key importance. In [29], we propose a conservative scheme, based on finite volume methods on an adaptive grid, which is able to render out the early steps of the reaction as well as the later chain reactions.

6.4. Liver organ modelling

Participants: Noémie Boissier, Dirk Drasdo, Géraldine Cellière, Adrian Friebel, Group Heinzle [Univ. Saarbruecken, Germany], Group Hengstler [IfADo, Germany], Stefan Hoehme, Tim Johann, Group Klingmueller [German Cancer Center, Heidelberg], Johannes Neitsch, Group Reo [Inria Paris - Rocquencourt], Paul Van Liedekerke, Eric Vibert [Hopital Paul Brousse], Yi Yin, Group Zerial [Max-Planck Inst. for Molecular Genetics, Dresden, Germany], Groups Iflow, Notox, Vln.

6.4.1. Ammonia detoxification after drug-induced damage.

The model for ammonia detoxification after drug-induced damage (see above) identified a systematic deviation between data and results that would be expected from the current standard model for ammonia detoxification in healthy liver [0] ([17], [6]) (see also comments/editorials in [0]). The findings triggered a series of new experiments identifying reversibility of the glutamate-dehydrogenase reaction in hepatocytes, and in blood (Ghallab et. al., subm.). It could be shown in an animal model that the newly recognized reactions can be therapeutically used to significantly reduce the concentration of toxic ammonia after drug-induced damage. (Work in close collaboration with partners of the project VLN (BMBF, Germany) and EU-NOTOX.

6.4.2. Systematic analysis strategies permitting quantitative conclusions in systems medicine and biology.

Based on the examples from liver regeneration after drug-induced damage [57] [17]) systematic iterative strategies can be inferred to enable identification of mechanisms underlying complex processes in spatial temporal tissue organisation and organ functioning. These use an iterative application of a pipeline of imaging, image analysis and modeling, quantitative models by parameterization of model components by measurable parameters for which the physiological ranges are known, and systematic simulated parameter sensitivity analyses [7].

[0] Haeussinger D., Eur. J. Biochem, 1983; Gebhardt R and Mecke, D. EMBO J 1983
[0] Wierling, C. Hepatology, 60(6) 2014; and: Widera, A., EXCLI Journal, 13, 2014

MYCENAE Project-Team

6. New Results

6.1. Highlights of the Year

- Picture of the Conference poster of the 2014 SIAM annual meeting (July 7-11, Chicago, USA), adapted from [7]
- Invitation to organize the mini symposium "The stochastic brain" at the Stochastic Processes and Applications Conference (Jul 28-Aug1, Buenos-Aires, Argentina)
- Selection of the NeuroMathMod project in the framework of the Sorbonne Université Emergence 2014 call

6.2. Numerical and theoretical studies of slow-fast systems with complex oscillations

6.2.1. A multiple time scale coupling of piecewise linear oscillators: Application to a neuroendocrine system

Participants: Frédérique Clément, Mathieu Desroches, Soledad Fernández García, Maciej Krupa.

We have analyzed a four dimensional slow-fast piecewise linear system consisting of two coupled oscillators [32]. Each oscillator is a continuous slow-fast piecewise linear system with three zones of linearity. The coupling is one-way, that is, one subsystem evolves independently and is forcing the other subsystem. We have analyzed not only the qualitative behavior, but also quantitative aspects such as the period, frequency and amplitude of the oscillations. The system is used to reproduce all the features endowed in a former smooth model and reproduce the secretion pattern of the hypothalamic neurohormone GnRH along the ovarian cycle in different species.

6.2.2. Border collision bifurcations of stroboscopic maps in periodically driven spiking models

Participants: Frédérique Clément, Albert Granados Corsellas, Maciej Krupa.

In [21], we have considered a general nonautonomous hybrid system based on the integrate-and-fire model, widely used as simplified version of neuronal models and other types of excitable systems. Our assumptions are that the system is monotonic, possesses an attracting subthreshold equilibrium point, and is forced by means of a periodic pulsatile (square wave) function. In contrast to classical methods, in our approach we use the stroboscopic map (time-T return map) instead of the so-called firing map. It becomes a discontinuous map potentially defined in an infinite number of partitions. By applying theory for piecewise-smooth systems, we avoid relying on particular computations, and we develop a novel approach that can be easily extended to systems with other topologies (expansive dynamics) and higher dimensions. More precisely, we have rigorously studied the bifurcation structure in the two-dimensional parameter space formed by the amplitude of the pulse and the ratio between T and the duration of the pulse (duty cycle). We show that it is covered by regions of existence of periodic orbits given by period adding structures. The period adding structures completely describe not only all the possible spiking asymptotic dynamics but also the behavior of the firing rate, which is a devil's staircase as a function of the parameters.

6.2.3. Interpreting frequency responses to dose-conserved pulsatile input signals in simple cell signaling motifs

Participants: Richard Bertram, Patrick Fletcher, Joël Tabak [Florida State University], Frédérique Clément, Alexandre Vidal.

Many hormones are released in pulsatile patterns. This pattern can be modified, for instance by changing pulse frequency, to encode relevant physiological information. Often other properties of the pulse pattern will also change with frequency. How do signaling pathways of cells targeted by these hormones respond to different input patterns? We have asked if a given dose of hormone can induce different outputs from the target system, depending on how this dose is distributed in time [20]. We have used simple mathematical models of feedforward signaling motifs to understand how the properties of the target system give rise to preferences in input pulse pattern. We frame these problems in terms of frequency responses to pulsatile inputs, where the amplitude or duration of the pulses is varied along with frequency to conserve input dose. We have found that nonlinearity in the steady state input-output function of the system predicts the optimal input pattern. It does so by selecting an optimal input signal amplitude. Our results predict the behavior of common signaling motifs such as receptor binding with dimerization, and protein phosphorylation. The findings have implications for experiments aimed at studying the frequency response to pulsatile inputs, as well as for understanding how pulsatile patterns drive biological responses via feedforward signaling pathways.

6.2.4. Mixed-mode oscillations due to a singular Hopf bifurcation in a forest pest model

Participants: Morten Brøns [Technical University of Denmark], Mathieu Desroches, Maciej Krupa.

We have revisited a three-dimensional model of forest pest where MMOs play an important role [17]. In this model, young trees are distinguished from old trees, and the pest feeds on old trees. The pest grows on a fast scale, the young trees on an intermediate scale, and the old trees on a slow scale. We have established that the main organizing center for the shape and oscillatory patterns of the solutions is not a folded-node singularity, which does exist in the system, but rather a singular Hopf bifurcation. A combination of a singular Hopf bifurcation and a weak return mechanism, characterized by a very small change in one of the variables, determines the features of the mixed-mode oscillations. Period-doubling and saddle-node bifurcations lead to closed families (called isolas) of periodic solutions in a bifurcation corresponding to a singular Hopf bifurcation.

6.2.5. On the Dynamics of the adenylate energy system: homeorhesis versus homeostasis

Participants: Jesús M Cortés, Ildefonso M. de La Fuente, Iker Malaina, Luis Martínez, Edelmira Valero [University of Bilbao], Serafim Rodrigues [Plymouth University], Mathieu Desroches.

We have developed and analyzed a new model of the ATP-ADP-AMP biochemical system in order to understand some of the functional elements involved in the cellular energy status [18]. In this model based on a delayed differential system, the enzymatic rate equations and all the physiological kinetic parameters have been explicitly considered and experimentally tested in vitro. Our central hypothesis is that cells are characterized by changing energy dynamics (homeorhesis). The results have shown that the adenylate energy charge (AEC) presents stable transitions between steady states and periodic oscillations and, in agreement with experimental data these oscillations range within the narrow AEC window. Furthermore, the model shows sustained oscillations in the Gibbs free energy and in the total nucleotide pool.

6.2.6. Adaptative algorithms for the simulation of slow-fast coupled oscillators in networks

Participants: Frédérique Clément, Marie Postel, Alexandre Vidal.

The numerical simulation of a slow fast system is usually performed using an explicit scheme with an adaptive time step, in order to preserve the numerical accuracy during the fast dynamic events. In the case of large sized networks of coupled slow-fast systems, one need to use the same very small time step for all components of the network, since the integration is performed simultaneously on the whole network. We have proposed a new algorithm based on a dynamic split of the network components, in the framework of symplectic integrators [40], and applied it to a model describing the intracellular calcium oscillations in a network of embryonic GnRH neurons [9]. At each time step, the systems currently in the fast dynamic parts, are identified from their distance to the fast manifold. These components are accordingly integrated using a small time step, while a larger time step is used for the remaining of the network (cf poster abstract in the CANUM 2014 conference). Although the CPU time saving is proportional to the time constant ratio between the slow and fast dynamics, it hardly compensates the drop in the convergence order as the size of the network increases.

6.3. Non conservative transport equations for cell population dynamics

6.3.1. 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. In [16], we tackle the difficulties arising when applying a Multiresolution method to a transport equation with discontinuous fluxes modeling localized mitosis. 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.

6.3.2. Calibration of a multiscale model for cell dynamics

Participants: Benjamin Aymard, Frédérique Clément, Marie Postel, Kim Long Tran.

In the framework of the PhD of Benjamin Aymard and the master training of Kim Long Tran, we have tackled the issue of the numerical calibration of our multiscale model of cell populations in ovarian follicles, in collaboration with Danielle Monniaux (INRA Tours). The strategy has consisted in designing quantitative specifications from the available biological knowledge, most of which fall within the field of cell population kinetics (e.g. growth fraction, mitotic index ...), and translating them into constraints on the model parameters, as well as in performing a detailed a priori analysis of the properties of the mathematical functions entering the model equations. Using visualization approaches appropriate both for following the trajectory of a given ovarian follicle with time and comparing the follicles together, we have confronted the model outputs on different levels (from the local cell density to the overall cell number) to the corresponding specifications. We have been able to reproduce instances of the selection process occurring within a cohort of terminally growing follicles. To enable one to do systematic explorations of the model behavior in different parameter configurations associated with either physiological (e.g. species-specific ovulation number) or pathological situations (dysovulation), we have undertaken a reduction approach inspired from [41]. We have generalized these results by relaxing some simplifying assumptions to account for some important features of the original model as the distinction between different phases in the cell division cycle.

6.4. Macroscopic limits of stochastic neural networks and neural fields

6.4.1. Pulsatile localized dynamics in delayed neural-field equations in arbitrary dimension

Participants: Jonathan Touboul, Grégory Faye [EHESS].

Neural field equations are integro-differential systems describing the macroscopic activity of spatially extended pieces of cortex. In such cortical assemblies, the propagation of information and the transmission machinery induce communication delays, due to the transport of information (propagation delays) and to the synaptic machinery (constant delays). We have investigated the role of these delays on the formation of structured spatiotemporal patterns for these systems in arbitrary dimensions [19]. We have focused on localized activity, either induced by the presence of a localized stimulus (pulses) or by transitions between two levels of activity (fronts). Linear stability analysis allows to reveal the existence of Hopf bifurcation curves induced by the delays, along different modes that may be symmetric or asymmetric. We show that instabilities strongly depend on the dimension, and in particular may exhibit transversal instabilities along invariant directions. These instabilities yield pulsatile localized activity, and depending on the symmetry of the destabilized modes, either produce spatiotemporal breathing or sloshing patterns.

6.4.2. Limits and dynamics of randomly connected neuronal networks

Participants: Cristóbal Quiñinao [CIRB], Jonathan Touboul.

Networks of the brain are composed of a very large number of neurons connected through a random graph and interacting after random delays that both depend on the anatomical distance between cells. In order to comprehend the role of these random architectures on the dynamics of such networks, we have analyzed the mesoscopic and macroscopic limits of networks with random correlated connectivity weights and delays [35]. We have addressed both averaged and quenched limits, and shown propagation of chaos and convergence to a complex integral McKean-Vlasov equations with distributed delays. We have then instantiated a completely solvable model illustrating the role of such random architectures in the emerging macroscopic activity. We have particularly focused on the role of connectivity levels in the emergence of periodic solutions.

6.4.3. The propagation of chaos in neural fields

Participant: Jonathan Touboul.

We have considered the problem of the limit of bio-inspired spatially extended neuronal networks including an infinite number of neuronal types (space locations), with space-dependent propagation delays modeling neural fields [24]. The propagation of chaos property is proved in this setting under mild assumptions on the neuronal dynamics, valid for most models used in neuroscience, in a mesoscopic limit, the neural-field limit, in which we can resolve the quite fine structure of the neuron activity in space and where averaging effects occur. The mean-field equations obtained are of a new type: they take the form of well-posed infinite-dimensional delayed integro-differential equations with a nonlocal mean-field term and a singular spatio-temporal Brownian motion. We have also shown how these intricate equations can be used in practice to uncover mathematically the precise mesoscopic dynamics of the neural field in a particular model where the mean-field equations exactly reduce to deterministic nonlinear delayed integro-differential equations.

6.4.4. Spatially extended networks with singular multi-scale connectivity patterns

Participant: Jonathan Touboul.

In [24], we took care of a number of technical difficulties arising in the description of large-scale systems that are spatially extended. The organization of neurons in space (within cortical columns) and their interactions (fully connected networks) were relatively far from what is known of the anatomy of neuronal networks. In [25], we have further taken into account the fine and macroscopic structure of the cortex, which is a very large network characterized by a complex connectivity including at least two scales. On the microscopic scale, the interconnections are non-specific and very dense, while macroscopic connectivity patterns connecting different regions of the brain at larger scale are extremely sparse. This motivates to analyze the behavior of networks with multiscale coupling, in which a neuron is connected to its $v(N)$ nearest-neighbors where $v(N) = o(N)$, and in which the probability of macroscopic connection between two neurons vanishes. These are called singular multi-scale connectivity patterns. We have introduced a class of such networks and derived their continuum limit. We show convergence in law and propagation of chaos in the thermodynamic limit. The limit equation obtained is an intricate non-local McKean-Vlasov equation with delays which is universal with respect to the type of micro-circuits and macro-circuits involved.

6.4.5. Index Distribution of the Ginibre Ensemble

Participants: Romain Allez [Stastlab, Cambridge University], Gilles Wainrib [ENS], Jonathan Touboul.

Complex systems, and in particular random neural networks, are often described by randomly interacting dynamical systems with no specific symmetry. In that context, characterizing the number of relevant directions necessitates fine estimates on the Ginibre ensemble. We have computed analytically the probability distribution of the number of eigenvalues N_R with modulus greater than R (the index) of a large $N \times N$ random matrix in the real or complex Ginibre ensemble [15]. We have shown that the fraction $N_R/N = p$ has a distribution scaling as $exp(-\beta N^2 \psi_R(p))$ with $\beta = 1$ (respectively $\beta = 1/2$) for the complex (resp. real) Ginibre ensemble. For any $p \in [0, 1]$, the equilibrium spectral densities as well as the rate function $\psi_R(p)$ are explicitly derived. This function displays a third order phase transition at the critical (minimum) value $p_R^* = 1 - R^2$, associated to a phase transition of the Coulomb gas. We have deduced that, in the central regime, the fluctuations of the index N_R around its typical value $p_R^* N$ scale as $N^{1/3}$.

6.4.6. The heterogeneous gas with singular interaction: Generalized circular law and heterogeneous renormalized energy

Participants: Luis-Carlos Garcia Del Molino, Khashayar Pakdaman [Institut Jacques Monod], Jonathan Touboul.

We have introduced and analyzed d dimensional Coulomb gases with random charge distribution and general external confining potential [23]. Our long term motivation is to understand the spectrum of random matrices with non identical distributions, for instance with independent elements with distinct statistics. We have shown that these gases satisfy a large deviation principle. The analysis of the minima of the rate function (which is the leading term of the energy) reveals that at equilibrium, the particle distribution is a generalized circular law (i.e. with spherical support but non-necessarily uniform distribution). In the classical electrostatic external potential, there are infinitely many minimizers of the rate function. The most likely macroscopic configuration is a disordered distribution in which particles are uniformly distributed (for $d = 2$, the circular law), and charges are independent of the positions of the particles. General charge-dependent confining potentials unfold this degenerate situation: in contrast, the particle density is not uniform, and particles spontaneously organize according to their charge. In that picture the classical electrostatic potential appears as a transition at which order is lost. Sub-leading terms of the energy are derived: we show that these are related to an operator, generalizing the Coulomb renormalized energy, which incorporates the heterogeneous nature of the charges. This heterogeneous renormalized energy informs us about the microscopic arrangements of the particles, which are non-standard, strongly depending on the charges, and include progressive and irregular lattices.

POMDAPI Project-Team

5. New Results

5.1. A posteriori error estimates

Participant: Martin Vohralík.

In [2], we have been able to derive an a posteriori error estimate for the numerical approximation of the two-phase flow problem. This is a cornerstone model problem for porous media, describing the flow of two immiscible and incompressible fluids. We take into account the capillary pressure, whence the model features such difficulties as coupling of partial differential equations with algebraic constraints, strong nonlinearities, degeneracy (disappearance of the diffusion term), advection dominance and consequent forming of sharp evolving fronts, or highly nonlinear and very badly conditioned systems of algebraic equations. Our analysis covers a large class of spatial discretizations in a unified setting, with fully implicit time stepping. We also show how the different error components, namely the spatial discretization error, the temporal discretization error, the linearization error, and the algebraic solver error can be distinguished and estimated separately. This gives rise to efficient adaptive stopping criteria, enabling to spare many useless iterations. The practical impact of our results is that even for this complicated model problem, the overall error committed in a numerical approximation can be fully controlled and, moreover, the simulation time can be reduced by factors typically of an order of magnitude. This result has then been extended in [4] to the compositional model of multiphase Darcy flow, where an arbitrary number of phases can be present, and where each phase can be composed of several components. Later, in [12], still a possible dependence on the temperature has been added. The last two references also contain convincing numerical illustrations on real-life reservoir engineering examples.

5.2. Optimization

Participants: Jean Charles Gilbert, Émilie Joannopoulos, Cédric Josz.

5.2.1. Polynomial optimization

A polynomial optimization problem (POP) consists in minimizing a multivariate real polynomial on a set K defined by polynomial inequalities and equalities. In its full generality it is a non-convex, multi-extremal, difficult global optimization problem. More than a decade ago, J. B. Lasserre proposed to solve a POP by a hierarchy of convex semidefinite programming (SDP) relaxations of increasing size and precision. Each problem in the hierarchy has a primal SDP formulation (a relaxation of a moment expression of the POP) and a dual SDP formulation (a sum-of-squares polynomial relaxation of the POP). In [18], we show that there is no duality gap between each primal and dual SDP problem in Lasserre's hierarchy, provided one of the constraints in the description of set K is a ball constraint. Our proof uses elementary results on SDP duality and it does not assume that K has a strictly feasible point.

5.2.2. Convex quadratic optimization

Convex quadratic optimization deals with problems consisting in minimizing a convex quadratic function on a polyhedron. In [3], we analyzed the behavior of the augmented Lagrangian algorithm when it deals with an *infeasible* convex quadratic optimization problem; this situation is important to master in order to be able to solve correctly the QPs that are generated by the SQP (or Newton-like) algorithm to solve a nonlinear optimization problem, QPs whose feasibility is not guaranteed. It is shown that the algorithm finds a point that, on the one hand, satisfies the constraints shifted by the smallest possible shift that makes them feasible and, on the other hand, minimizes the objective on the corresponding shifted constrained set. The speed of convergence to such a point is globally linear, with a rate that is inversely proportional to the augmentation parameter. This suggests a rule for determining the augmentation parameter that aims at controlling the speed of convergence of the shifted constraint norm to zero; this rule has the advantage of generating bounded augmentation parameters even when the problem is infeasible. The approach has also been implemented in the pieces of software OQLA and QPALM during the ADT MINOQS (see section 4.2 and [16], [14], [15]).

REO Project-Team

6. New Results

6.1. Highlights of the Year

- Jimmy Mullaert was awarded the best poster prize at the conference Canum 2014.
- Jessica Oakes was awarded a University of California Presidential Postdoctoral Fellowship.
- Jessica Oakes won a young investigator award at the "4th International Conference on Engineering Frontiers in Pediatric and Congenital Heart Disease".

6.2. Mathematical and numerical analysis of fluid-structure interaction problems

Participants: Benoit Fabrèges, Miguel Ángel Fernández Varela, Mikel Landajuela Larma, Jimmy Mullaert, Marina Vidrascu.

- In [54] we introduce two new classes of numerical methods for the solution of incompressible fluid/thin-walled structure interaction problems with unfitted meshes. The semi-implicit or explicit nature of the splitting in time is dictated by the order in which the spatial and time discretizations are performed. Stability and optimal accuracy are achieved without restrictive CFL conditions or correction iterations. Results presented by M. Landajuela at the 11th World Congress on Computational Mechanics (WCCM XI), July 20-25, 2014, Barcelona (Spain).
- In [47] 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. The time splitting combines a projection method in the fluid with a specific Robin-Neumann treatment of the interface coupling. A priori energy estimates guaranteeing unconditional stability are established for some of the schemes. The accuracy and performance of the methods proposed is illustrated by a thorough numerical study.
- We have performed an a priori error analysis for the generalized Robin-Neumann explicit coupling schemes introduced in [30]. The analysis confirms the $\mathcal{O}(\tau^{2^{r-1}}/h^{\frac{1}{2}})$ error perturbation anticipated by the numerical evidence of [30]. Another fundamental result of this work is that the h-non-uniformity of the splitting error is not a consequence of the mass-lumping approximation (which simply dictates the explicit or semi-implicit nature of the coupling scheme). The analysis indicates that the genesis of the $\mathcal{O}(h^{-\frac{1}{2}})$ is the non-uniformity of discrete viscoelastic operator, which is a consequence of thick-walled character of the solid. These results have been reported in [48] and presented by M.A. Fernández at the 11th World Congress on Computational Mechanics (WCCM XI), July 20-25, 2014, Barcelona (Spain).
- We consider the extension of the Nitsche-XFEM method to fluid-structure interaction problems involving a thin-walled elastic structure (Lagrangian formalism) immersed in an incompressible fluid (Eulerian formalism). The fluid domain is discretized with an unstructured mesh not fitted to the solid mid- surface mesh. Weak and strong discontinuities across the interface are allowed for the velocity and pressure, respectively. The kinematic/kinetic fluid-solid coupling is enforced consistently using a variant of Nitsche's method involving cut elements. Robustness with respect to arbitrary interface/element intersections is guaranteed through a ghost penalty stabilization. Different coupling schemes, either fully implicit or loosely coupled, are proposed. Several numerical examples, involving static and moving interfaces, illustrate the performance of the methods. A paper in collaboration with F. Alauzet (project-team Gamma3) is under preparation. Results presented by B. Fabrèges at the 11th World Congress on Computational Mechanics (WCCM XI), July 20-25, 2014, Barcelona (Spain).

6.3. Numerical methods for biological flows

Participants: Grégory Arbia, Benoit Fabrèges, Miguel Ángel Fernández Varela, Justine Fouchet-Incaux, Jean-Frédéric Gerbeau, Céline Grandmont, Sanjay Pant, Saverio Smaldone, Marc Thiriet, Irène Vignon-Clementel.

- In [19] We consider the problem of estimating the stiffness of an artery wall using a data assimilation method applied to a 3D fluid-structure interaction (FSI) model. We briefly present the FSI model, the data assimilation procedure based on a reduced order Unscented Kalman filter, and the segmentation algorithm. We then present two examples of the procedure using real data. First, we estimate the stiffness distribution of a silicon rubber tube from image data. Second, we present the estimation of aortic wall stiffness from real clinical data.
- In [29], 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 [55] we investigate the stability of numerical schemes that are classically used in the simulation of airflows and blood flows. The geometrical complexity of the networks in which air/blood flows leads to a classical decomposition of two areas: a truncated 3D geometry corresponding to the largest contribution of the domain, and a 0D part connected to the 3D part, modelling air/blood flows in smaller airways/vessels. The resulting Navier-Stokes system in the 3D truncated part may involve non-local boundary conditions, deriving from a mechanical model. For various 3D/0D coupled models, different discretization processes are presented and analyzed in terms of numerical stability, highlighting strong differences according to the regimes that are considered. In particular, two main stability issues are investigated: first the coupling between the 3D and the 0D part for which implicit or explicit strategies are studied and, second, the question of estimating the amount of kinetic energy entering the 3D domain because of the artificial boundaries. The second issue has been also the subject of a review [31].
- In [31] we deal with numerical simulations of incompressible Navier-Stokes equations in truncated domain. In this context, the formulation of these equations has to be selected carefully in order to guarantee that their associated artificial boundary conditions are relevant for the considered problem. In this paper, we review some of the formulations proposed in the literature, and their associated boundary conditions. Some numerical results linked to each formulation are also presented. We compare different schemes, giving successful computations as well as problematic ones, in order to better understand the difference between these schemes and their behaviours dealing with systems involving Neumann boundary conditions. We also review two stabilization methods which aim at suppressing the instabilities linked to these natural boundary conditions.
- In [40], we propose a framework for Windkessel parameter estimation in a 0D representation of the 3D fluid-flow domain. Parameters are estimated from uncertain measurements through a sequential approach, and the 0D representation is iteratively improved through 3D-CFD simulations. The application of generalized sensitivity functions to assess parameter correlation and to ascertain the measurement set needed to avoid identifiability problems is also presented through representative test cases. This method, which is capable of handling non-simultaneous measurements, is demonstrated and validated for a patient-specific case of aortic coarctation.
- In [17] we perform the first patient-specific pulmonary hemodynamics 3D-0D modeling before single ventricle stage 2 surgery. 0D parameters are automatically tuned to match flow and pressure clinical measurements that are not taken where 3D boundary conditions need to be specified. This work on six patients demonstrates how simulations can help to check the coherence of clinical data or provide insights to clinicians that are otherwise difficult to measure, such as in the presence of kinks.

- In [25] we study a case of post single ventricle stage 2 surgery with the three following aims: (i) to show how to build a patient-specific model describing the hemodynamics in the presence of collaterals, using patient-specific clinical data collected at different times; (ii) to use this model to perform virtual collateral occlusion for quantitative hemodynamics prediction; and (iii) to compare predicted hemodynamics with post-operative measurements.

6.4. Numerical methods for cardiac electrophysiology

Participants: Muriel Boulakia, Jean-Frédéric Gerbeau, Damiano Lombardi, Elisa Schenone.

- In [33], a reduced-order method based on Approximated Lax Pairs (ALP) is applied to the integration of electrophysiology models. These are often high- dimensional parametric equation systems, challenging from a model reduction stand- point. The method is tested on two and three dimensional test-cases, of increasing complexity. The solutions are compared to the ones obtained by a finite element. The reduced-order simulation of pseudo-electrocardiograms based on ALP is proposed in the last part.
- In [21], we address the question of the discretization of Stochastic Partial Differential Equations (SPDE) for excitable media. Working with SPDE driven by colored noise, we consider a numerical scheme based on finite differences in time (Euler-Maruyama) and finite elements in space. Motivated by biological considerations, we study numerically the emergence of reentrant patterns in excitable systems such as the Barkley or Mitchell-Schaeffer models.

6.5. Lung and respiration modeling

Participants: Laurent Boudin, Muriel Boulakia, Céline Grandmont, Jessica Oakes, Ayman Moussa, Irène Vignon-Clementel.

- In [20], we consider the non-reactive fully elastic Boltzmann equation for mixtures. We deduce that, under the standard diffusive scaling, its limit for vanishing Mach and Knudsen numbers is the Maxwell-Stefan model for a multicomponent gaseous mixture.
- In [49], we first deal with the modelling and the discretization of an aerosol evolving in the air, in the respiration framework, within a domain which can be fixed or moving. We also investigate basic numerical properties of the numerical code which was developped, and also focus on the influence of the aerosol on the airflow.
- In [38], the aim of the study was to determine susceptibility differences between healthy and emphysematous rats exposed to airborne particles. To do this, we performed animal exposure experimenters and measured particle deposition concentrations with Magnetic Resonance Imaging. We showed that overall deposition was significantly higher in the elastase-treated rats compared to the healthy ones, suggesting enhanced susceptibility to airborne particles in diseased lungs. Current work aims at integrating such experimental data into modeling [39] and compare numerical simulations with experiments. To extend particle modeling to expiration, a 1D particle transport model is under development [44].

  While it is known that the retention of fine particles is less in microgravity (uG) compared to normal gravity (1G) levels, it was unknown the spatial relationship of deposited particles. In [26], rats were exposed to 1 micron diameter particles on the NASA uG airplane and compared to rats exposed in 1G. We found that the ratio of deposited particles in the central airways compared to the peripheral ones, was significantly less in the uG than in 1G, indicating enhanced deposition in the periphery. This data suggests that toxicology effects of exposure to Moon dust may not be insignificant.
- In [51], we establish stability estimates for the unique continuation property of the nonstationary Stokes problem. These estimates hold without prescribing boundary conditions and are of logarithmic type. They are obtained thanks to Carleman estimates for parabolic and elliptic equations. Then, these estimates are applied to an inverse problem where we want to identify a Robin coefficient defined on some part of the boundary from measurements available on another part of the boundary.

6.6. Miscellaneous

Participants: Jean-Frédéric Gerbeau, Damiano Lombardi, Marina Vidrascu.

- in [32] a reduced-order model algorithm, called ALP, is proposed to solve nonlinear evolution partial differential equations. It is based on approximations of generalized Lax pairs. Contrary to other reduced-order methods, like Proper Orthogonal Decomposition, the basis on which the solution is searched for evolves in time according to a dynamics specific to the problem. It is therefore well-suited to solving problems with progressive front or wave propagation. Another difference with other reduced-order methods is that it is not based on an off-line / on-line strategy. Numerical examples are shown for the linear advection, KdV and FKPP equations, in one and two dimensions.
- in [41] we propose a direct method for computing modal coupling coefficients - due to geometrically nonlinear effects - for thin shells vibrating at large amplitude and discretized by a finite element (FE) procedure. These coupling coefficients arise when considering a discrete expansion of the unknown displacement onto the eigenmodes of the linear operator. The evolution problem is thus projected onto the eigenmodes basis and expressed as an assembly of oscillators with quadratic and cubic nonlinearities. The nonlinear coupling coefficients are directly derived from the finite element formulation, with specificities pertaining to the shell elements considered, namely, here elements of the "Mixed Interpolation of Tensorial Components" family (MITC). Therefore, the computation of coupling coefficients, combined with an adequate selection of the significant eigenmodes, allows the derivation of effective reduced-order models for computing - with a continuation procedure - the stable and unstable vibratory states of any vibrating shell, up to large amplitudes. The procedure is illustrated on a hyperbolic paraboloid panel. Bifurcation diagrams in free and forced vibrations are obtained. Comparisons with direct time simulations of the full FE model are given. Finally, the computed coefficients are used for a maximal reduction based on asymptotic nonlinear normal modes (NNMs), and we find that the most important part of the dynamics can be predicted with a single oscillator equation.
- in [53] we deal with the following data assimilation problem: construct an analytical approximation of a numerical constitutive law in three-dimensional nonlinear elasticity. More precisely we are concerned with a micro-macro model for rubber as the one proposed in [36]. Macroscopic quantities of interest such as the Piola-Kirchhoff stress tensor can be approximated for any value of the strain gradient by numerically solving a nonlinear PDE. This procedure is however computationally demanding. Hence, although conceptually satisfactory, this physically-based model is of no direct practical use. We aim to circumvent this difficulty by proposing a numerical strategy to reconstruct from in silico experiments an accurate analytical proxy for the micro-macro constitutive law.

SISYPHE Project-Team

5. New Results

5.1. Fault detection and localization in networks of transmission lines

Participants: Mohamed Oumri, Michel Sorine.

Some results have been obtained in collaboration with Florent Loete (LGEP) and Qinghua Zhang (Inria, I4S):
- *Experimental validation of the inverse scattering method for distributed characteristic impedance estimation.* Our theoretic results and numerical simulations have shown the ability of inverse scattering-based methods to diagnose soft faults in electric cables, in particular, faults implying smooth spatial variations of cable characteristic parameters. We have obtained laboratory experiments confirming the ability of the inverse scattering method for retrieving spatially distributed characteristic impedance from reflectometry measurements. Various smooth or stepped spatial variations of characteristic impedance profiles have been tested. The tested electric cables are CAN unshielded twisted pairs used in trucks and coaxial cables [37].

- *Diagnosis of networks using tagged electric lines.* A new electromagnetic marking method of transmission lines has been proposed for diagnosis of electric networks when conditions of uniqueness of the solution are not fulfilled (e.g. in case of symmetries): small non-interfering characteristic defaults are added to the lines and used as tags. A patent application has been submitted [36].

A new application of our monitoring technique has been explored in collaboration with EDF and a first result has been obtained:
- *Monitoring of post-tensioned ducts or water content in concrete walls with embedded transmission lines.* We have presented an electromagnetic method of diagnosis based on frequency domain reflectometry (FDR) associated with our inversion algorithm, ISTL (Inverse Scattering for Transmission Lines). ISTL allows one to estimate the spatial profile of the electrical impedance of the line from the FDR measurements. Experimental results on two mockups of external post-tensioned ducts with filling defects show the feasibility of the method. We will try to show the similarities between auscultation external post-tensioned ducts and measurement of water content by TDR probes (Time Domain Reflectometry) [34].

- *Fault diagnosis of wired electric networks by reflectometry.* A first extension to Baum-Liu-Tesche equations has been proposed in [31].

5.2. Cardiovascular signal processing and applications

Participants: Lisa Guigue, Claire Médigue, Michel Sorine, Serge Steer.

See the Sofware section 4.1 for a description of tools developed for *Cardiovascular Waves Analysis.*

5.3. Glycemic control in ICUs

Participant: Michel Sorine.

The results of statistical analysis of the data gathered during the large clinical trial CGAO-REA have been published in [14]: “Tight computerized versus conventional glucose control in the ICU: a randomized controlled trial”. Despite the increase in the incidence of severe hypoglycemia in our experimental group, based on the absence of difference in mortality between patients on tight computerized glucose control and those on less stringent glucose control without computerized decision support systems (CDSS), this study could pave the way for future randomized controlled trials assessing new generation CDSSs allowing the safe implementation of blood glucose control in the ICU that take into account the complexity of glucose control throughout the ICU stay and the variability of individualized insulin needs. Some new objectives for computer aided glycemic control in ICUs have been proposed in [32]. An article proposing a more detailed statistical analysis of the severe hypoglycemic events has been submitted.

5.4. Modeling and optimizing patient pathways in hospital

Participants: James Leifer, Michel Sorine.

External scientific collaboration with:
- Niccolo Curatolo, Directeur des opérations, Hôpitaux universitaires Paris-Sud, Assistance publique-Hôpitaux de Paris (AP-HP);
- Dr Maurice Raphaël, Chef de service, Urgences adultes, Hôpital Bicêtre, AP-HP;
- Dr Christophe Vincent-Cassy, Responsable des systèmes informatiques des urgences, AP-HP;
- Lucie Gaillardot-Roussel, Ingénieur en organisation, AP-HP;
- Dr Paul Jarvis, Senior consultant doctor in emergency medicine, Calderdale and Huddersfield National Health Service Foundation Trust, UK.

In 2014, we began a case study of the emergency department (ED) at Bicêtre Hospital, a large ED handling 50,000 patient visits per year, which is amongst the top 10 by volume and by annual volume growth for EDs in the Paris region.

Rather than presume the appropriateness of a predetermined scientific formalism, our strategy was to allow the application to frame a series of questions in order to lead us to experiment with several potential scientific tools at the present "low risk, high uncertainty" phase of investigation:

- *Top-down modeling:* Can we capture the expert knowledge of doctors and nurses as to the pathways followed by their patients by transforming this knowledge into a series of "use case" rules borrowed from the techniques of software specification? Can these rules by transformed into an executable model using business process modeling languages and tools (Orc, YAWL, ...) for simulating the complex parallel composition of man-machine processes in a hospital setting?

- *Bottom-up modeling:* How can the hospital be instrumented for cheaply and accurately capturing its real activity (movement of people and machines, delays, errors, ...) and tuning the parameters of the model? Can we intercept HL7 messages (a standardized electronic message format for medical data) and/or access raw time-stamped database entries to use machine learning techniques (particularly process mining) to extract from the running hospital the graphs representing the actual sequence of care events in order to get rapid feedback about the most heavily used and most often delayed path segments?

- *Underlying cost semantics:* Can we formalize in process calculi (for example, a variation of pi calculus) the "micro internal economy" of costs exchanged inside the hospital to quantify the economic performance of each patient pathway?

- *Offline experimentation and optimization:* Can potential optimization to the model be explored offline in a sort of "serious game" to allow non-intrusive experimentation with different strategies for eliminating bottlenecks, increasing flow rates, decreasing costs, etc.?

- *Data visualization for medical personnel:* Given that the medical personnel themselves are best suited to fixing the daily frictional time losses that most are resigned to accept as "part of the job", how can the model be presented in a visually lucid manner to render the previously "invisible" aspects of the hospital's organization visible?

- *Online real-time control:* Can the feedback loop be completed and the model be used to directly provide real-time visual feedback to the hospital personnel to enable them to measure their systemic progress (or systemic unintended consequences) of their localized optimizations?

ALPINES Project-Team

6. New Results

6.1. Highlights of the Year

We have released a version of FreeFem++ (v 3.33) which introduces new and important features related to high performance computing:

- Interface with PETSc library
- Interface with HPDDM (see above)
- improved interface with the parallel direct solver MUMPS

This release enables, for the first time, end-users to run the very same code on computers ranging from laptops to clusters and even large scale computers with thousands of computing nodes

6.2. Communication avoiding algorithms for dense linear algebra

Our group continues to work on algorithms for dense linear algebra operations that minimize communication. During this year we focused on improving the performance of communication avoiding QR factorization as well as designing algorithms that reduce communication on multilevel hierarhical platforms.

In [17] we focus on the QR factorization. The Tall-Skinny QR (TSQR) algorithm is more communication efficient than the standard Householder algorithm for QR decomposition of matrices with many more rows than columns. However, TSQR produces a different representation of the orthogonal factor and therefore requires more software development to support the new representation. Further, implicitly applying the orthogonal factor to the trailing matrix in the context of factoring a square matrix is more complicated and costly than with the Householder representation. We show how to perform TSQR and then reconstruct the Householder vector representation with the same asymptotic communication efficiency and little extra computational cost. We demonstrate the high performance and numerical stability of this algorithm both theoretically and empirically. The new Householder reconstruction algorithm allows us to design more efficient parallel QR algorithms, with significantly lower latency cost compared to Householder QR and lower bandwidth and latency costs compared with Communication-Avoiding QR (CAQR) algorithm. As a result, our final parallel QR algorithm outperforms ScaLAPACK and Elemental implementations of Householder QR and our implementation of CAQR on the Hopper Cray XE6 NERSC system.

In [18] we focus on performance predictions of multilevel communication optimal LU and QR factorizations on hierarchical platforms. This study focuses on the performance of two classical dense linear algebra algorithms, the LU and the QR factorizations, on multilevel hierarchical platforms. We first introduce a new model called Hierarchical Cluster Platform (HCP), encapsulating the characteristics of such platforms. The focus is set on reducing the communication requirements of studied algorithms at each level of the hierarchy. Lower bounds on communications are therefore extended with respect to the HCP model. We then introduce multilevel LU and QR algorithms tailored for those platforms, and provide a detailed performance analysis. We also provide a set of numerical experiments and performance predictions demonstrating the need for such algorithms on large platforms.

6.3. Enlarged Krylov methods

Krylov subspace methods are among the most practical and popular iterative methods today. They are polynomial iterative methods that aim to solve systems of linear equations ($Ax = b$) by finding a sequence of vectors $x_1, x_2, x_3, x_4, ..., x_k$ that minimizes some measure of error over the corresponding spaces $x_0 + \mathcal{K}_i(A, r_0), \quad i = 1, ..., k$ where $\mathcal{K}_i(A, r_0) = span\{r_0, Ar_0, A^2r_0, ..., A^{i-1}r_0\}$ is the Krylov subspace of dimension i, x_0 is the initial iterate, and r_0 is the initial residual. These methods are governed by Blas1 and

Blas2 operations as dot products and sparse matrix vector multiplications. Parallelizing dot products is constrained by communication since the performed computation is negligible. If the dot products are performed by one processor, then there is a need for a communication before and after the computation. In both cases, communication is a bottleneck. In [21] we introduce a new approach for reducing communication in Krylov subspace methods that consists of enlarging the Krylov subspace by a maximum of t vectors per iteration, based on the domain decomposition of the graph of A. The obtained enlarged Krylov subspace $\mathcal{K}_{t,k}(A, r_0)$ is a superset of the Krylov subspace $\mathcal{K}_k(A, r_0)$, $\mathcal{K}_k(A, r_0) \subset \mathcal{K}_{t,k+1}(A, r_0)$. Thus it is possible to search for the solution of the system $Ax = b$ in $\mathcal{K}_{t,k}(A, r_0)$ instead of $\mathcal{K}_k(A, r_0)$. Moreover, we show that the enlarged Krylov projection subspace methods lead to faster convergence in terms of iterations and parallelizable algorithms with less communication, with respect to Krylov methods.

6.4. Algebraic preconditioners

Our work focused on the design of robust algebraic preconditioners and domain decomposition methods to accelerate the convergence of iterative methods.

In [8] we introduce the block filtering decomposition, a new preconditioning technique that is suitable for matrices arising from the discretization of a system of PDEs on unstructured grids. The preconditioner satisfies a so-called filtering property, which ensures that the input matrix is identical with the preconditioner on a given filtering vector. This vector is chosen to alleviate the effect of low frequency modes on convergence and so decrease or eliminate the plateau which is often observed in the convergence of iterative methods. In particular, the paper presents a general approach that allows to ensure that the filtering condition is satisfied in a matrix decomposition. The input matrix can have an arbitrary sparse structure. Hence, it can be reordered using nested dissection, to allow a parallel computation of the preconditioner and of the iterative process. We present experimental results that demonstrate the efficiency of the proposed preconditioner on a set of matrices arising from the discretization of partial differential equations on two-dimensional and three-dimensional grids. We also show that the numerical efficiency of the preconditioner does not suffer from the reordering of the unknowns for the matrices in our test set, which can have highly heterogeneous and anisotropic coefficients.

In [9] we discuss the usage of overlapping techniques for improving the convergence of preconditioners based on incomplete factorizations. To enable parallelism, these preconditioners are usually applied after the input matrix is permuted into nested bordered block diagonal form. We use k-way partitioning with vertex separator (KPVS) to recursively partitions the corresponding graph of the input matrix into k subgraphs using a subset of its vertices called separators. In the case where $k = 2$, it is called nested dissection. The overlapping technique is then based on algebraically extending the associated subdomains of these subgraphs and their corresponding separators obtained from KPVS by their direct neighbours. This approach is known to accelerate the convergence of domain decomposition methods, where the input matrix is partitioned into a number of independent subdomains using k-way graph partitioning, a different graph decomposition technique. We discuss the effect of the overlapping technique on the convergence of two classes of preconditioners, based on nested factorization and block incomplete LDU factorization.

In [22] we introduce LORASC, a robust algebraic preconditioner for solving sparse linear systems of equations involving symmetric and positive definite matrices. The graph of the input matrix is partitioned by using k-way partitioning with vertex separators into N disjoint domains and a separator formed by the vertices connecting the N domains. The obtained permuted matrix has a block arrow structure. The preconditioner relies on the Cholesky factorization of the first N diagonal blocks and on approximating the Schur complement corresponding to the separator block. The approximation of the Schur complement involves the factorization of the last diagonal block and a low rank correction obtained by solving a generalized eigenvalue problem or a randomized algorithm. The preconditioner can be build and applied in parallel. Numerical results on a set of matrices arising from the discretization by the finite element method of linear elasticity models illustrate the robusteness and the efficiency of our preconditioner.

The Helmholtz equation governing wave propagation and scattering phenomena is difficult to solve numerically. Its discretization with piecewise linear finite elements results in typically large linear systems of equations. The inherently parallel domain decomposition methods constitute hence a promising class of precondi-

tioners. An essential element of these methods is a good coarse space. Here, the Helmholtz equation presents a particular challenge, as even slight deviations from the optimal choice can be devastating.

In [5], we present a coarse space that is based on local eigenproblems involving the Dirichlet-to-Neumann operator. Our construction is completely automatic, ensuring good convergence rates without the need for parameter tuning. Moreover, it naturally respects local variations in the wave number and is hence suited also for heterogeneous Helmholtz problems. The resulting method is parallel by design and its efficiency is demonstrated on 2D homogeneous and heterogeneous numerical examples.

Coarse spaces are instrumental in obtaining scalability for domain decomposition methods for partial differential equations (PDEs). However, it is known that most popular choices of coarse spaces perform rather weakly in the presence of heterogeneities in the PDE coefficients, especially for systems of PDEs. In [12], we introduce in a variational setting a new coarse space that is robust even when there are such heterogeneities. We achieve this by solving local generalized eigenvalue problems in the overlaps of subdomains that isolate the terms responsible for slow convergence. We prove a general theoretical result that rigorously establishes the robustness of the new coarse space and give some numerical examples on two and three dimensional heterogeneous PDEs and systems of PDEs that confirm this property.

Multiphase, compositional porous media flow models lead to the solution of highly heterogeneous systems of Partial Differential Equations (PDEs). In [7] , we focus on overlapping Schwarz type methods on parallel computers and on multiscale methods. We recall a coarse space that is robust even when there are such heterogeneities. The two-level domain decomposition approach is compared to multiscale methods.

In [16], we investigate two-level preconditioners on the extended linear system arising from the domain decomposition method. The additive Schwarz method is used as a smoother, and the coarse grid space is constructed by using the Ritz vectors obtained in the Arnoldi process. The coarse grid space can be improved adaptively as the Ritz vectors become a better approximation of the eigenvectors. Numerical tests on the model problem demonstrate the efficiency.

6.5. New results related to FreeFem++

In [10], we propose an efficient algorithm for the numerical approximation of metrics, used for anisotropic mesh adaptation on triangular meshes with finite element computations. We derive the metrics from interpolation error estimates expressed in terms of higher order derivatives, for the $P-k$-Lagrange finite element, $k>1$. Numerical examples of mesh adaptation done using metrics computed with our Algorithm, and derived from higher order derivatives as error estimates, show that we obtain the right directions of anisotropy.

In [2], we consider a system of two reaction-dispersion equations with non constant parameters. Both equations are coupled through the boundary conditions. We propose a mixed variational formulation that leads to a non symmetric saddle-point problem. We prove its well-posedness. Then, we develop a stabilized mixed finite element discretization of this problem and establish optimal a priori error estimates.

In [15], we consider a model of soil water and nutrient transport with plant root uptake. The geometry of the plant root system is explicitly taken into account in the soil model. We first describe our modeling approach. Then, we introduce an adaptive mesh refinement procedure enabling us to accurately capture the geometry of the root system and small-scale phenomena in the rhizosphere. Finally, we present a domain decomposition technique for solving the problems arising from the soil model as well as some numerical results.

6.6. Auto adaptative algorithms

In [29], we develop an adaptive version of the inexact Uzawa algorithm applied to finite element discretizations of the linear Stokes problem. We base our developments on an equilibrated flux a posteriori error estimate distinguishing the different error components, namely the discretization error component, the inner algebraic solver error component, and the outer Uzawa iteration error component. On each outer Uzawa and inner linear algebraic solver iteration, we prove that our estimate gives a guaranteed upper bound on the total error, as well as a polynomial-degree-robust local efficiency. Our adaptive inexact algorithm stops the outer Uzawa iteration

and the inner linear algebraic solver iteration when the Uzawa error component, respectively the algebraic solver error component, do not have a significant influence on the total error. The developed framework covers all standard conforming and conforming stabilized finite element methods. The implementation into the FreeFem++ programming language is invoked and two numerical examples showcase the performance of our adaptive strategy.

6.7. Spectrum for a small inclusion of negative material

We studied a spectral problem $(\mathcal{P}^\delta)$ for a diffusion like equation in a 3D domain Ω. The main originality here lies in the presence of a parameter σ^δ, whose sign changes on Ω, in the principal part of the operator we consider. More precisely, σ^δ is positive on Ω except in a small inclusion of size $\delta > 0$. Because of the sign-change of σ^δ, for all $\delta > 0$ the spectrum of $(\mathcal{P}^\delta)$ consists of two sequences converging to $+\infty$ and $-\infty$. However, at the limit $\delta = 0$, the small inclusion vanishes so that there should only remain positive spectrum for $(\mathcal{P}^\delta)$. What happens to the negative spectrum? In this paper, we prove that the positive spectrum of $(\mathcal{P}^\delta)$ tends to the spectrum of the problem without the small inclusion. On the other hand, we establish that each negative eigenvalue of $(\mathcal{P}^\delta)$ behaves like $\delta^{-2}\mu$ for some constant $\mu < 0$. We also show that the eigenvectors associated with the negative eigenvalues are localized around the small inclusion. We end the article providing 2D numerical experiments illustrating these results.

6.8. Stability of electromagnetic cavities perturbed by small perfectly conducting inclusions

We consider an electromagnetic wave propagation problem in harmonic regime in a bounded cavity, in the case where the medium of propagation contains small perfectly conducting inclusions. We prove that the solution to this problem depends continuously on the data in a uniform manner with respect to the size of the inclusions.

6.9. Integral equations for acoustic scattering by partially impenetrable composite objects

We study direct first-kind boundary integral equations arising from transmission problems for the Helmholtz equation with piecewise constant coefficients and Dirichlet boundary conditions imposed on a closed surface. We identify necessary and sufficient conditions for the occurrence of so-called spurious resonances, that is, the failure of the boundary integral equations to possess unique solutions.

Following [A. Buffa and R. Hiptmair, Numer Math, 100, 1–19 (2005)] we propose a modified version of the boundary integral equations that is immune to spurious resonances. Via a gap construction it will serve as the basis for a universally well-posed stabilized global multi-trace formulation that generalizes the method of [X. Claeys and R. Hiptmair, Commun Pure and Appl Math, 66, 1163–1201 (2013)] to situations with Dirichlet boundary conditions.

6.10. Application domain: data analysis in astrophysics

One of the application domain on which our algorithms are validated is data analysis in astrophysics. Estimation of the sky signal from sequences of time order data is one of the key steps in the Cosmic Microwave Background (CMB) data analysis, commonly referred to as the map-making problem. Some of the most popular and general methods proposed for this problem involve solving generalised least squares (GLS) equations with non-diagonal noise weights given by a block-diagonal matrix with Toeplitz blocks. In [14] we study new map-making solvers potentially suitable for applications to the largest, anticipated data sets. They are based on iterative conjugate gradient (CG) approaches enhanced with novel, parallel, two-level preconditioners (2lvl-PCG). We apply the proposed solvers to examples of simulated, non-polarised and polarised CMB observations and a set of idealised scanning strategies with a sky coverage ranging from nearly a full sky down to small sky patches. We discuss in detail their implementation for massively parallel

computational platforms and their performance for a broad range of parameters characterising the simulated data sets. We find that our best new solver can outperform carefully optimised, standard solvers as used today, by as much as a factor of 5 in terms of the convergence rate and a factor of 4 in terms of the time to solution, and does so without increasing significantly the memory consumption or the volume of inter-processor communication. The performance of the new algorithms is also found to be more stable, robust and less dependent on specific characteristics of the analysed data set.We therefore conclude that the proposed approaches are well suited to address successfully challenges posed by new and forthcoming CMB data sets.

Spherical Harmonic Transforms (SHT) are at the heart of many scientific and practical applications ranging from climate modelling to cosmological observations. In many of these areas new, cutting-edge science goals have been recently proposed requiring simulations and analyses of experimental or observational data at very high resolutions and of unprecedented volumes. Both these aspects pose formidable challenge for the currently existing implementations of the transforms.

In [13] we describe parallel algorithms for computing SHT with two variants of intra-node parallelism appropriate for novel supercomputer architectures, multi-core processors and Graphic Processing Units (GPU). It also discusses their performance, alone and embedded within a top-level, MPI-based parallelisation layer ported from the S^2HAT library, in terms of their accuracy, overall efficiency and scalability. We show that our inverse SHT run on GeForce 400 Series GPUs equipped with latest CUDA architecture ("Fermi") outperforms the state of the art implementation for a multi-core processor executed on a current Intel Core i7-2600K. Furthermore, we show that an MPI/CUDA version of the inverse transform run on a cluster of 128 Nvidia Tesla S1070 is as much as 3 times faster than the hybrid MPI/OpenMP version executed on the same number of quad-core processors Intel Nehalem for problem sizes motivated by our target applications. Performance of the direct transforms is however found to be at the best comparable in these cases. We discuss in detail the algorithmic solutions devised for the major steps involved in the transforms calculation, emphasising those with a major impact on their overall performance, and elucidates the sources of the dichotomy between the direct and the inverse operations.

DYOGENE Project-Team

6. New Results

6.1. Highlights of the Year

- F. Baccelli received 2014 IEEE Communications Society Stephen O. Rice Prize in the Field of Communications Theory:

  http://www.comsoc.org/about/memberprograms/comsoc-awards/rice.
- F. Baccelli received 2014 IEEE Communications Society Leonard G. Abraham Prize in the Field of Communications Systems:

  http://www.comsoc.org/about/memberprograms/comsoc-awards/abraham.
- F. Baccelli received ACM Sigmetrics Achievement Award 2014:

  http://www.sigmetrics.org/achievementaward-2014.shtml.
- F. Simatos received 2014 ACM SIGMETRICS Rising Star Researcher Award:

  http://www.sigmetrics.org/risingstar-2014.shtml.
- P. Brémaud published a book "Fourier Analysis and Stochastic Processes". Series: Universitext. Springer, Sept. 2014 - 385 pages.
- PhD student C. Rovetta received best tool paper award at Valuetools 2014 for the paper [18].

6.2. On Spatial Point Processes with Uniform Births and Deaths by Random Connection

With I. Norros (VTT Finland) and F. Mathieu (Bell Labs France), F. Baccelli has continued the line of thought on the geometry of Peer-to-Peer systems that was initiated in their Infocom 13 paper. This type of dynamics leads to a class of spatial birth and death process of the Euclidean space where the birth rate is constant and the death rate of a given point is the shot noise created at its location by the other points of the current configuration for some response function f. An equivalent view point is that each pair of points of the configuration establishes a random connection at an exponential time determined by f, which results in the death of one of the two points. The research concentrated on space-motion invariant processes of this type. Under some natural conditions on f, one can construct the unique time-stationary regime of this class of point processes by a coupling argument. The birth and death structure can then be used to establish a hierarchy of balance integral relations between the factorial moment measures. One can also show that the time-stationary point process exhibits a certain kind of repulsion between its points that is called f-repulsion.

These results were published in [29].

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

With A. Giovanidis, IMT, F. Baccelli has studied a cooperation model where the positions of base stations 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 base stations are involved. If two, these cooperate by exchange of user data and reduced channel information (channel phase, second neighbour interference) with conferencing over some backhaul link. The total user transmission power is split between them and a common message is encoded, which is coherently transmitted by the stations. The decision for a user to choose service with or without cooperation is directed by a family of geometric policies. The suggested policies further control the shape of coverage contours in favor of cell-edge areas. Analytic expressions based on stochastic geometry are derived for the coverage probability in the network. Their numerical evaluation shows benefits from cooperation, which are enhanced when Dirty Paper Coding is applied to eliminate the second neighbour interference.

These results were published in [7].

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

With C. Singh (IIT), F. Baccelli and B. Blaszczyszyn worked on combining adaptive protocol design, utility maximization and stochastic geometry. The focus was on a spatial adaptation of Aloha within the framework of ad hoc networks. Quasi-static networks are considered, in which mobiles learn the local topology and incorporate this information to adapt their medium access probability (MAP) selection to their local environment. The cases where nodes cooperate in a distributed way to maximize the global throughput or to achieve either proportional fair or max-min fair medium access were considered. The proportionally fair sharing case 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. It was shown 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. 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.

These results were published in [16].

6.5. Quality of Real-Time Streaming in Wireless Cellular Networks - Stochastic Modeling and Analysis

We present a new stochastic service model with capacity sharing and interruptions, appropriate for the evaluation of the quality of real-time streaming (e.g. mobile TV) in wireless cellular networks [2]. It takes into account multi-class Markovian process of call arrivals (to capture different radio channel conditions, requested streaming bit-rates and call-durations) and allows for a general resource allocation policy saying which users are temporarily denied the requested fixed streaming bit-rates (put in outage) due to resource constraints. We develop general expressions for the performance characteristics of this model, including the mean outage duration and the mean number of outage incidents for a typical user of a given class, involving only the steady-state of the traffic demand. We propose also a natural class of least-effort-served-first resource allocation policies, which cope with optimality and fairness issues known in wireless networks, and whose performance metric! s can be easily calculated using Fourier analysis of Poisson variables. We specify and use our model to analyze the quality of real time streaming in 3GPP Long Term Evolution (LTE) cellular networks. Our results can be used for the dimensioning of these networks.

6.6. On Comparison of Clustering Properties of Point Processes

In [3], we propose a new comparison tool for spatial homogeneity of point processes, based on the joint examination of void probabilities and factorial moment measures. We prove that determinantal and permanental processes, as well as, more generally, negatively and positively associated point processes are comparable in this sense to the Poisson point process of the same mean measure. We provide some motivating results on percolation and coverage processes, and preview further ones on other stochastic geometric models, such as minimal spanning forests, Lilypond growth models, and random simplicial complexes, showing that the new tool is relevant for a systemic approach to the study of macroscopic properties of non-Poisson point processes. This new comparison is also implied by the directionally convex ordering of point processes, which has already been shown to be relevant to the comparison of the spatial homogeneity of point processes. For this latter ordering, using a notion of lattice perturbation, we provide a large monotone spectrum of comparable point processes, ranging from periodic grids to Cox processes, and encompassing Poisson point processes as well. They are intended to serve as a platform for further theoretical and numerical studies of clustering, as

well as simple models of random point patterns to be used in applications where neither complete regularity nor the total independence property are realistic assumptions.

6.7. SINR in Wireless Networks and the Two-Parameter Poisson-Dirichlet Process

Stochastic geometry models of wireless networks based on Poisson point processes are increasingly being developed with a focus on studying various signal-to-interference-plus-noise ratio (SINR) values. In [9], we show that the SINR values experienced by a typical user with respect to different base stations of a Poissonian cellular network are related to a specific instance of the so-called two-parameter Poisson-Dirichlet process. This process has many interesting properties as well as applications in various fields. We give examples of several results proved for this process that are of immediate or potential interest in the development of analytic tools for cellular networks. Some of them simplify or are akin to certain results that are being developed in the network literature. By doing this we hope to motivate further research and use of Poisson-Dirichlet processes in this new setting.

6.8. How User Throughput Depends on the Traffic Demand in Large Cellular Networks

In [17], 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 law 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 st! ate 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.

6.9. Pioneers of Influence Propagation in Social Networks

In [20], we present a diffusion model developed by enriching the generalized random graph (a.k.a. configuration model), motivated by the phenomenon of viral marketing in social networks. The main results on this model are rigorously proved in [3], and in this paper we focus on applications. Specifically, we consider random networks having Poisson and Power Law degree distributions where the nodes are assumed to have varying attitudes towards influence propagation, which we encode in the model by their transmitter degrees. We link a condition involving total degree and transmitter degree distributions to the effectiveness of a marketing campaign. This suggests a novel approach to decision-making by a firm in the context of viral marketing which does not depend on the detailed information of the network structure.

6.10. QoS and Network Performance Estimation in Heterogeneous Cellular Networks Validated by Real-Field Measurements

Mobile network operators observe a significant disparity of quality of service (QoS) and network performance metrics, such as the mean user throughput, the mean number of users and the cell load, over different network base stations. The principal reason being the fact that real networks are never perfectly hexagonal, base stations are subject to different radio conditions, and may have different engineering parameters. In [21], we propose a model that takes into account these network irregularities in a probabilistic manner, in particular assuming Poisson spatial location of base stations, lognormal shadowing and random transmission powers. Performance of base stations is modeled by spatial processor sharing queues, which are made dependent of each other via a system of load equations. In order to validate our approach, we estimate all the model parameters from the data collected in a commercial network, solve it and compare the spatial variability of the QoS and performance metrics! in the model to the real network performance metrics. Considering two scenarios: downtown of a big city and a mid-size city, we show that our model predicts well the network performance.

6.11. Clustering Comparison of Point Processes with Applications to Random Geometric Models

In [27], we review some examples, methods, and recent results involving comparison of clustering properties of point processes. Our approach is founded on some basic observations allowing us to consider void probabilities and moment measures as two complementary tools for capturing clustering phenomena in point processes. As might be expected, smaller values of these characteristics indicate less clustering. Also, various global and local functionals of random geometric models driven by point processes admit more or less explicit bounds involving void probabilities and moment measures, thus aiding the study of impact of clustering of the underlying point process. When stronger tools are needed, directional convex ordering of point processes happens to be an appropriate choice, as well as the notion of (positive or negative) association, when comparison to the Poisson point process is considered. We explain the relations between these tools and provide examples of point processes admitting them. Furthermore, we sketch some recent results obtained using the aforementioned comparison tools, regarding percolation and coverage properties of the germ-grain model, the SINR model, subgraph counts in random geometric graphs, and more generally, U-statistics of point processes. We also mention some results on Betti numbers for Čech and Vietoris-Rips random complexes generated by stationary point processes. A general observation is that many of the results derived previously for the Poisson point process generalise to some "sub-Poisson" processes, defined as those clustering less than the Poisson process in the sense of void probabilities and moment measures, negative association or dcx-ordering.

6.12. 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 [11], 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.

6.13. How Clustering Affects Epidemic in Random Networks

Motivated by the analysis of social networks, we study a model of random networks that has both a given degree distribution and a tunable clustering coefficient. We consider two types of growth processes on these graphs: diffusion and symmetric threshold model. The diffusion process is inspired from epidemic models. It is characterized by an infection probability, each neighbor transmitting the epidemic independently. In the symmetric threshold process, the interactions are still local but the propagation rule is governed by a threshold (that might vary among the different nodes). An interesting example of symmetric threshold process is the contagion process, which is inspired by a simple coordination game played on the network. Both types of processes have been used to model spread of new ideas, technologies, viruses or worms and results have been obtained for random graphs with no clustering. In [6], we are able to analyze the impact of clustering on the growth processes. While clustering inhibits the diffusion process, its impact for the contagion process is more subtle and depends on the connectivity of the graph: in a low connectivity regime, clustering also inhibits the contagion, while in a high connectivity regime, clustering favors the appearance of global cascades but reduces their size. For both diffusion and symmetric threshold models, we characterize conditions under which global cascades are possible and compute their size explicitly, as a function of the degree distribution and the clustering coefficient. Our results are applied to regular or power-law graphs with exponential cutoff and shed new light on the impact of clustering.

6.14. Edge Label Inference in Generalized Stochastic Block Model: From Spectral Theory to Impossibility Results

The classical setting of community detection consists of networks exhibiting a clustered structure. To more accurately model real systems we consider a class of networks (i) whose edges may carry labels and (ii) which may lack a clustered structure. Specifically we assume that nodes possess latent attributes drawn from a general compact space and edges between two nodes are randomly generated and labeled according to some unknown distribution as a function of their latent attributes. Our goal is then to infer the edge label distributions from a partially observed network. In [22], we propose a computationally efficient spectral algorithm and show it allows for asymptotically correct inference when the average node degree could be as low as logarithmic in the total number of nodes. Conversely, if the average node degree is below a specific constant threshold, we show that no algorithm can achieve better inference than guessing without using the observations. As a byproduct of our analysis, we show that our model provides a general procedure to construct random graph models with a spectrum asymptotic to a pre-specified eigenvalue distribution such as a power-law distribution.

6.15. Balanced Graph Edge Partition

Balanced edge partition has emerged as a new approach to partition an input graph data for the purpose of scaling out parallel computations, which is of interest for several modern data analytics computation platforms, including platforms for iterative computations, machine learning problems, and graph databases. This new approach stands in a stark contrast to the traditional approach of balanced vertex partition, where for given number of partitions, the problem is to minimize the number of edges cut subject to balancing the vertex cardinality of partitions.

In [19], we first characterize the expected costs of vertex and edge partitions with and without aggregation of messages, for the commonly deployed policy of placing a vertex or an edge uniformly at random to one of the partitions. We then obtain the first approximation algorithms for the balanced edge-partition problem which for the case of no aggregation matches the best known approximation ratio for the balanced vertex-partition problem, and show that this remains to hold for the case with aggregation up to factor that is equal to the maximum in-degree of a vertex. We report results of an extensive empirical evaluation on a set of real-world graphs, which quantifies the benefits of edge- vs. vertex-partition, and demonstrates efficiency of natural greedy online assignments for the balanced edge-partition problem with and with no aggregation.

6.16. Streaming, Memory-limited Algorithms for Community Detection

In [23], we consider sparse networks consisting of a finite number of non-overlapping communities, i.e. disjoint clusters, so that there is higher density within clusters than across clusters. Both the intra- and inter-cluster edge densities vanish when the size of the graph grows large, making the cluster reconstruction problem nosier and hence difficult to solve. We are interested in scenarios where the network size is very large, so that the adjacency matrix of the graph is hard to manipulate and store. The data stream model in which columns of the adjacency matrix are revealed sequentially constitutes a natural framework in this setting. For this model, we develop two novel clustering algorithms that extract the clusters asymptotically accurately. The first algorithm is *offline*, as it needs to store and keep the assignments of nodes to clusters, and requires a memory that scales linearly with the network size. The second algorithm is *online*, as it may classify a node when the corresponding column is revealed and then discard this information. This algorithm requires a memory growing sub-linearly with the network size. To construct these efficient streaming memory-limited clustering algorithms, we first address the problem of clustering with partial information, where only a small proportion of the columns of the adjacency matrix is observed and develop, for this setting, a new spectral algorithm which is of independent interest.

6.17. State Space Collapse for Critical Multistage Epidemics

We study a multistage epidemic model which generalizes the SIR model and where infected individuals go through $K > 0$ stages of the epidemic before being removed. An infected individual in stage k may infect a susceptible individual, who directly goes to stage k of the epidemic; or it may go to the next stage $k + 1$ of the epidemic. For this model, we identify the critical regime in which we establish diffusion approximations. Surprisingly, the limiting diffusion exhibits an unusual form of state space collapse which we analyze in detail.

6.18. Perfect Sampling for Closed Queueing Networks

In [4], we investigate coupling from the past (CFTP) algorithms for closed queueing networks. The stationary distribution has a product form only in a very limited number of particular cases when queue capacity is finite, and numerical algorithms are intractable due to the cardinality of the state space. Moreover, closed networks do not exhibit any monotonic property enabling efficient CFTP. We derive a bounding chain for the CFTP algorithm for closed queueing networks. This bounding chain is based on a compact representation of sets of states that enables exact sampling from the stationary distribution without considering all initial conditions in the CFTP. The coupling time of the bounding chain is almost surely finite, and numerical experiments show that it is close to the coupling time of the exact chain.

In [18], we present Clones, a Matlab toolbox for exact sampling from the stationary distribution of a closed queueing net-work with finite capacities. This toolbox is based on recent results using a compact representation of sets of states that enables exact sampling from the stationary distribu-tion without considering all initial conditions in the coupling from the past (CFTP) scheme. This representation reduces the complexity of the one-step transition in the CFTP al-gorithm to O(KM 2), where K is the number of queues and M the total number of customers; while the cardinality of the state space is exponential in the number of queues. In this paper, we focus on the algorithmic and implementation issues. We propose a new representation, that leads to one-step transition complexity of the CFTP algorithm that is in O(KM). We provide a detailed description of our matrix-based implementation. The toolbox can be downloaded at http://www.di.ens.fr/~rovetta/Clones.

6.19. Individual Risk in Mean-Field Control Models for Decentralized Control, with Application to Automated Demand Response

Flexibility of energy consumption can be harnessed for the purposes of ancillary services in a large power grid. In prior work by the authors a randomized control architecture is introduced for individual loads for this purpose. In examples it is shown that the control architecture can be designed so that control of the loads is easy at the grid level: Tracking of a balancing authority reference signal is possible, while ensuring

that the quality of service (QoS) for each load is acceptable on average. The analysis was based on a mean field limit (as the number of loads approaches infinity), combined with an LTI-system approximation of the aggregate nonlinear model. In [15], we examine in depth the issue of individual risk in these systems. The main contributions of the paper are of two kinds: Risk is modeled and quantified: (i) The average performance is not an adequate measure of success. It is found empirically that a histogram of QoS is approximately Gaussian, and consequently each load will eventually receive poor service. (ii) The variance can be estimated from a refinement of the LTI model that includes a white-noise disturbance; variance is a function of the randomized policy, as well as the power spectral density of the reference signal. Additional local control can eliminate risk: (iii) The histogram of QoS is truncated through this local control, so that strict bounds on service quality are guaranteed. (iv) This has insignificant impact on the grid-level performance, beyond a modest reduction in capacity of ancillary service.

6.20. Passive Dynamics in Mean Field Control

Mean-field models are a popular tool in a variety of fields. They provide an understanding of the impact of interactions among a large number of particles or people or other "self-interested agents", and are an increasingly popular tool in distributed control. In [14], we consider a particular randomized distributed control architecture introduced in our own recent work. In numerical results it was found that the associated mean-field model had attractive properties for purposes of control. In particular, when viewed as an input-output system, its linearization was found to be minimum phase. In this paper we take a closer look at the control model. The results are summarized as follows: (i) The Markov Decision Process framework of Todorov is extended to continuous time models, in which the "control cost" is based on relative entropy. This is the basis of the construction of a family of controlled Markovian generators. (ii) A decentralized control architecture is proposed in which each agent evolves as a controlled Markov process. A central authority broadcasts a common control signal to each agent. The central authority chooses this signal based on an aggregate scalar output of the Markovian agents. (iii) Provided the control-free system is a reversible Markov process, the following identity holds for the linearization,

$$\text{Real}(G(j\omega)) = \text{PSD}_Y(\omega) \geq 0 \qquad \omega \in \mathbb{R}\,,,$$

where the right hand side denotes the power spectral density for the output of any one of the individual (control-free) Markov processes.

6.21. Optimization of Dynamic Matching Models

The bipartite matching model was born in the work of Gale and Shapley, who proposed the stable marriage problem in the 1960s. In [36], we consider a dynamic setting, modeled as a multi-class queueing network or MDP model. The goal is to compute a policy for the matching model that is optimal in the average cost sense. Computation of an optimal policy is not possible in general, but we obtain insight by considering relaxations. The main technical result is a form of "heavy traffic" asymptotic optimality. For a parameterized family of models in which the network load approaches capacity, a variant of the MaxWeight policy is approximately optimal, with bounded regret, even though the average cost grows without bound. Numerical results demonstrate that the policies introduced in this paper typically have much lower cost when compared to polices considered in prior work.

6.22. Stochastic Bounds with a Low Rank Decomposition

In [5], we investigate how we can bound a discrete time Markov chain (DTMC) by a stochastic matrix with a low rank decomposition. We show how the complexity of the analysis for steady-state and transient distributions can be simplified when we take into account the decomposition. Finally, we show how we can obtain a monotone stochastic upper bound with a low rank decomposition.

6.23. Generalizations of Bounds on the Index of Convergence to Weighted Digraphs

Sequences of maximum-weight walks of a growing length in weighted digraphs have many applications in manufacturing and transportation systems, as they encode important performance parameters. It is well-known that they eventually enter a periodic regime if the digraph is strongly connected. The length of their transient phase depends, in general, both on the size of digraph and on the magnitude of the weights. In this paper, we show that certain bounds on the transients of unweighted digraphs, such as the bounds of Wielandt, Dulmage-Mendelsohn, Schwarz, Kim, and Gregory-Kirkland-Pullman, remain true for critical nodes in weighted digraphs.

This work was done by Thomas Nowak together with Glenn Merlet from Aix-Marseille Unversité, Hans Schneider from the University of Wisconsin at Madison, and Sergeĭ Sergeev from the University of Birmingham. It was presented at the 53th IEEE Conference on Decision and Control and appeared in the journal Discrete Applied Mathematics.

6.24. Approximate Consensus in Highly Dynamic Networks: The Role of Averaging Algorithms

In this paper, we investigate the approximate consensus problem in highly dynamic networks in which topology may change continually and unpredictably. We prove that in both synchronous and partially synchronous systems, approximate consensus is solvable if and only if the communication graph in each round has a rooted spanning tree, i.e., there is a coordinator at each time. The striking point in this result is that the coordinator is not required to be unique and can change arbitrarily from round to round. Interestingly, the class of averaging algorithms which are memoryless and require no process identities entirely captures the solvability issue of approximate consensus in that the problem is solvable if and only if it can be solved using any averaging algorithm. Concerning the time complexity of averaging algorithms, we show that approximate consensus can be achieved with precision of ε in a coordinated network model in $O(n^{n+1}\log 1/\varepsilon)$ synchronous rounds, and in $O((\Delta n)^{n\Delta+1}\log 1/\varepsilon)$ rounds when the maximum round delay for a message to be delivered is Δ. We investigate various network models in which this exponential bound in the number of nodes reduces to a polynomial bound, and we prove that a general upper bound on the time complexity of averaging algorithms has to be exponential. We apply our results to networked systems with a fixed topology and classical benign fault models, and deduce both known and new results for approximate consensus in these systems. In particular, we show that for solving approximate consensus, a complete network can tolerate up to $2n-3$ arbitrarily located link faults at every round, in contrast with the impossibility result established by Santoro and Widmayer (STACS '89) showing that exact consensus is not solvable with $n-1$ link faults per round originating from the same node.

This work was done by Thomas Nowak together with Bernadette Charron-Bost from the CNRS and Matthias Függer from Vienna University of Technology. It is currently under submission.

6.25. Towards Binary Circuit Models That Faithfully Capture Physical Solvability

In contrast to analog models, binary circuit models are high-level abstractions that play an important role in assess-ing the correctness and performance characteristics of digital circuit designs: (i) modern circuit design relies on fast digital timing simulation tools and, hence, on binary-valued circuit models that faithfully model signal propagation, even throughout a complex design, and (ii) binary circuit models provide a level of abstraction that is amenable to formal correctness proofs. A mandatory feature of any such model is the ability to trace glitches and other short pulses precisely as they occur in physical circuits, as their presence may affect a circuit's correctness and its performance characteristics. Unfortunately, it was recently proved [Függer et al., ASYNC'13] that none of the existing binary-valued circuit models proposed so far, including the two most commonly used pure and inertial delay channels and any other bounded single-history channel, is realistic

in the following sense: For the simple Short-Pulse Filtration (SPF) problem, which is related to a circuit's ability to suppress a single glitch, they showed that every bounded single-history channel either contradicts the unsolvability of SPF in bounded time or the solvability of SPF in unbounded time in physical circuits, i.e., no existing model correctly captures physical solvability with respect to glitch propagation. We propose a binary circuit model, based on so-called in-volution channels, which do not suffer from this deficiency. In sharp contrast to what is possible with all the existing models, they allow to solve the SPF problem precisely when this is possible in physical circuits. To the best of our knowledge, our involution channel model is hence the very first binary circuit model that realistically models glitch propagation, which makes it a promising candidate for developing more accurate tools for simulation and formal verification of digital circuits.

This work was done by Thomas Nowak together with Matthias Függer, Robert Najvirt, and Ulrich Schmid from Vienna University of Technolgy. It will be presented at the conference DATE 2105.

6.26. Weak CSR Expansions and Transience Bounds in Max-Plus Algebra

This paper aims to unify and extend existing techniques for deriving upper bounds on the transient of max-plus matrix powers. To this aim, we introduce the concept of weak CSR expansions: $A^t = CS^tR \oplus B^t$. We observe that most of the known bounds (implicitly) take the maximum of (i) a bound for the weak CSR expansion to hold, which does not depend on the values of the entries of the matrix but only on its pattern, and (ii) a bound for the CStR term to dominate. To improve and analyze (i), we consider various cycle replacement techniques and show that some of the known bounds for indices and exponents of digraphs apply here. We also show how to make use of various parameters of digraphs. To improve and analyze (ii), we introduce three different kinds of weak CSR expansions. As a result, we obtain a collection of bounds, in general incomparable to one another, but better than the bounds found in the literature.

This work was done by Thomas Nowak together with Glenn Merlet from Aix-Marseille Unversité and Sergeĭ Sergeev from the University of Birmingham. It appeared in the journal Linear Algebra and its Applications.

6.27. An Overview of Transience Bounds in Max-Plus Algebra

This book chapter surveys and discusses upper bounds on the length of the transient phase of max-plus linear systems and sequences of max-plus matrix powers. In particular, It explains how to extend a result by Nachtigall to yield a new approach for proving such bounds and states an asymptotic tightness result by using an example given by Hartmann and Arguelles.

This work was done by Thomas Nowak together with Bernadette Charron-Bost from the CNRS. It appeared in the book "Tropical and Idempotent Mathematics and Applications" in the AMS's book series Contemporary Mathematics.

GANG Project-Team

5. New Results

5.1. Highlights of the Year

Pierre Fraigniaud has received the Prize for Innovation in Distributed Computing 2014.

5.2. Graph and Combinatorial Algorithms

5.2.1. Collision-Free Network Exploration

In the collision-free exploration model considered in [16], a set of mobile agents is placed at different nodes of a n-node network. The agents synchronously move along the network edges in a collision-free way, i.e., in no round may two agents occupy the same node. In each round, an agent may choose to stay at its currently occupied node or to move to one of its neighbors. An agent has no knowledge of the number and initial positions of other agents. We are looking for the shortest possible time required to complete the collision-free *network exploration*, i.e., to reach a configuration in which each agent is guaranteed to have visited all network nodes and has returned to its starting location.

In this work, we first considered the scenario when each mobile agent knows the map of the network, as well as its own initial position. We established a connection between the number of rounds required for collision-free exploration and the degree of the minimum-degree spanning tree of the graph. We provided tight (up to a constant factor) lower and upper bounds on the collision-free exploration time in general graphs, and the exact value of this parameter for trees. For our second scenario, in which the network is unknown to the agents, we proposed collision-free exploration strategies running in $O(n^2)$ rounds for tree networks and in $O(n^5 \log n)$ rounds for general networks.

5.2.2. Properties of Graph Search Procedures

In [4], we study the last vertex discovered by a graph search such as BFS or DFS. End-vertices of a given graph search may have some nice properties (as for example it is well known that the last vertex of Lexicographic Breadth First Search (LBFS) in a chordal graph is simplicial). Therefore it is interesting to consider if these vertices can be recognized in polynomial time or not. A graph search is a mechanism for systematically visiting the vertices of a graph. At each step of a graph search, the key point is the choice of the next vertex to be explored. Graph searches only differ by this selection mechanism during which a tie-break rule is used. In this paper we study how the choice of the tie-break rule can determine the complexity of the end-vertex problem for BFS or DFS. In particular we prove a counter-intuitive NP-completeness result for Breadth First Search, answering a question of D.G. Corneil, E. Köhler and J-M Lanlignel.

5.2.3. Matchings in Hypergraphs

A rainbow matching for (not necessarily distinct) sets $F_1, ...F_k$ of hypergraph edges is a matching consisting of k edges, one from each F_i. The aim of [3] is twofold—to put order in the multitude of conjectures that relate to this concept (some first presented here), and to prove partial results on one of the central conjectures settled by Ryser, Brualdi and Stein.

5.2.4. Common Intervals and Application to Genome Comparison

In [6], we show how to identify generalized common and conserved nested intervals. This is a bio-informatics papers, explaining how to compute more relaxed variants of common or of conserved intervals of two permutations, which has applications in genome comparison. It also presents some properties of the family of intervals, useful for storing them.

5.2.5. Graph Decomposition

In [10], we present a general framework for computing a large family of graph decomposition, the H-join. It generalizes some well know tools like modular decomposition or split decomposition. The paper explains how to compute it in polynomial time. A new canonical decomposition for sesquiprime graphs is also presented.

5.2.6. Combinatorial Optimization

Normal cone and subdifferential have been generalized through various continuous functions; in [8], we focus on a non separable Q-subdifferential version. Necessary and sufficient optimality conditions for unconstrained nonconvex problems are revisited accordingly. For inequality constrained problems, Q-subdifferential and the lagrangian multipliers, enhanced as continuous functions instead of scalars, allow us to derive new necessary and sufficient optimality conditions. In the same way, the Legendre-Fenchel conjugate is generalized into Q-conjugate and global optimality conditions are derived by Q-conjugate as well, leading to a tighter inequality.

5.3. Distributed Computing

5.3.1. Rendezvous

5.3.1.1. Rendezvous of Anonymous Agents in Trees

In [5], we study the so-called *rendezvous problem* in the mobile agent setting in graph environments. In the studied model, two identical (anonymous) mobile agents start from arbitrary nodes of an unknown tree and have to meet at some node. Agents move in synchronous rounds: in each round an agent can either stay at the current node or move to one of its neighbors. We consider deterministic algorithms for this rendezvous task. The main result of our research is a tight trade-off between the optimal time of completing rendezvous and the size of memory of the agents. For agents with k memory bits, we show that optimal rendezvous time is $\Theta(n + n^2/k)$ in n-node trees. More precisely, if $k \geq c \log n$, for some constant c, we design agents accomplishing rendezvous in arbitrary trees of size n (unknown to the agents) in time $O(n + n^2/k)$, starting with arbitrary delay. We also show that no pair of agents can accomplish rendezvous in time $o(n + n^2/k)$, even in the class of lines of known length and even with simultaneous start. Finally, we prove that at least logarithmic memory is necessary for rendezvous, even for agents starting simultaneously in a n-node line.

5.3.1.2. Rendezvous of Distance-Aware Mobile Agents in Unknown Graphs

In [17], we study the problem of rendezvous of two mobile agents starting at distinct locations in an unknown graph. The agents have distinct labels and walk in synchronous steps. However, the graph is unlabeled and the agents have no means of marking the nodes of the graph and cannot communicate with or see each other until they meet at a node. When the graph is very large, we would like the time to rendezvous to be independent of the graph size and to depend only on the initial distance between the agents and some local parameters such as the degree of the vertices, and the size of the agent's label. It is well known that even for simple graphs of degree Δ, the rendezvous time can be exponential in Δ in the worst case. In this study, we introduce a new version of the rendezvous problem where the agents are equipped with a device that measures its distance to the other agent after every step. We show that these *distance-aware* agents are able to rendezvous in any unknown graph, in time polynomial in all the local parameters such the degree of the nodes, the initial distance D and the size of the smaller of the two agent labels $l = \min(l_1, l_2)$. Our algorithm has a time complexity of $O(\Delta(D + \log l))$ and we show an almost matching lower bound of $\Omega(\Delta(D + \log l / \log \Delta))$ on the time complexity of any rendezvous algorithm in our scenario. Further, this lower bound extends existing lower bounds for the general rendezvous problem without distance awareness.

5.3.1.3. Rendezvous of Heterogeneous Mobile Agents in Edge-Weighted Networks

In [22], we study the deterministic rendezvous problem in which a pair of heterogeneous agents, differing in the time required to traverse particular edges of the graph, need to meet on an edge or node of the graph. Each of the agents knows the complete topology of the undirected graph and the initial positions of both of the agents. The agent also knows its own traversal times for all of the edges of the graph, but is unaware of the corresponding traversal times for the other agent. In this scenario, we study the time required by the agents

to meet, compared to the time T_{OPT} in the offline scenario in which the agents have complete knowledge of each others capabilities. When no additional assumptions are made, we show that rendezvous can be achieved after time $O(nT_{\mathrm{OPT}})$ in a n-node graph, and that this time is essentially the best possible in some cases. However, the rendezvous time can be reduced to $\Theta(T_{\mathrm{OPT}})$ when the agents are allowed to exchange $\Theta(n)$ bits of information at the start of the rendezvous process. We then show that under some natural assumption about the traversal times of edges, the hardness of the heterogeneous rendezvous problem can be substantially decreased, both in terms of time required for rendezvous without communication, and the communication complexity of achieving rendezvous in time $\Theta(T_{\mathrm{OPT}})$.

5.3.1.4. Rendezvous with Different Speeds

In [32] we introduce the study of the rendezvous problem in the context of agents having different speeds, and present tight and almost tight bounds for this problem, restricted to a ring topology.

5.3.2. Fair Synchronization

A non-blocking implementation of a concurrent object is an implementation that does not prevent concurrent accesses to the internal representation of the object, while guaranteeing the deadlock-freedom progress condition without using locks. Considering a failure free context, G. Taubenfeld has introduced (DISC 2013) a simple modular approach, captured under a new problem called the *fair synchronization* problem, to transform a non-blocking implementation into a starvation-free implementation satisfying a strong fairness requirement.

This approach is illustrated in [19] with the implementation of a concurrent stack. The spirit of the paper is mainly pedagogical. Its aim is not to introduce new concepts or algorithms, but to show that a powerful, simple, and modular transformation can provide concurrent objects with strong fairness properties.

In [20], we extend this approach in several directions. It first generalizes the fair synchronization problem to read/write asynchronous systems where any number of processes may crash. Then, it introduces a new failure detector and uses it to solve the fair synchronization problem when processes may crash. This failure detector, denoted QP (Quasi Perfect), is very close to, but strictly weaker than, the perfect failure detector. Last but not least, the paper shows that the proposed failure detector QP is optimal in the sense that the information on failures it provides to the processes can be extracted from any algorithm solving the fair synchronization problem in the presence of any number of process crash failures.

5.3.3. Wait Free with Advice

In [7], we motivate and propose a new way of thinking about failure detectors which allows us to define, quite surprisingly, what it means to solve a distributed task *wait-freeusing a failure detector*. In our model, the system is composed of *computation* processes that obtain inputs and are supposed to produce outputs and *synchronization* processes that are subject to failures and can query a failure detector.

Under the condition that *correct* synchronization processes take sufficiently many steps, they provide the computation processes with enough *advice* to solve the given task wait-free: every computation process outputs in a finite number of its own steps, regardless of the behavior of other computation processes.

Every task can thus be characterized by the *weakest* failure detector that allows for solving it, and we show that every such failure detector captures a form of set agreement. We then obtain a complete classification of tasks, including ones that evaded comprehensible characterization so far, such as renaming or weak symmetry breaking.

5.3.4. Adaptive Register Allocation

In [18], 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.

5.3.5. Leader Election

Considering the case of homonyms processes (some processes may share the same identifier) on a ring [21], we give a necessary and sufficient condition on the number of identifiers to enable leader election. We prove that if l is the number of identifiers then message-terminating election is possible if and only if l is greater than the greatest proper divisor of the ring size even if the processes do not know the ring size. If the ring size is known, we propose a process-terminating algorithm exchanging $O(n \log(n))$ messages that is optimal.

5.3.6. Concurrency and Fault-tolerance

In [15], we study the connections between self-stabilization and proof-labeling schemes. It follows from the definition of *silent* self-stabilization, and from the definition of *proof-labeling* scheme, that if there exists a silent self-stabilizing algorithm using ℓ-bit registers for solving a task T, then there exists a proof-labeling scheme for T using registers of at most ℓ bits. The first result in this paper is the converse to this statement. We show that if there exists a proof-labeling scheme for a task T, using ℓ-bit registers, then there exists a silent self-stabilizing algorithm using registers of at most $O(\ell + \log n)$ bits for solving T, where n is the number of processes in the system. Therefore, as far as memory space is concerned, the design of silent self-stabilizing algorithms essentially boils down to the design of compact proof-labeling schemes. The second result in this paper addresses time complexity. We show that, for every task T with k-bits output size in n-node networks, there exists a silent self-stabilizing algorithm solving T in $O(n)$ rounds, using registers of $O(n^2 + kn)$ bits. Therefore, as far as running time is concerned, *every* task has a silent self-stabilizing algorithm converging in a linear number of rounds.

In [27], we study the connections between, on the one hand, asynchrony and concurrency, and, on the other hand, the quality of the expected solution of a distributed algorithm. The state machine approach is a well-known technique for building distributed services requiring high performance and high availability, by replicating servers, and by coordinating client interactions with server replicas using consensus. Indulgent consensus algorithms exist for realistic eventually partially synchronous models, that never violate safety and guarantee liveness once the system becomes synchronous. Unavoidably, these algorithms may never terminate, even when no processor crashes, if the system never becomes synchronous. We propose a mechanism similar to state machine replication, called *RC-simulation*, that can always make progress, even if the system is never synchronous. Using RC-simulation, the quality of the service will adjust to the current level of asynchrony of the network — degrading when the system is very asynchronous, and improving when the system becomes more synchronous. RC-simulation generalizes the state machine approach in the following sense: when the system is asynchronous, the system behaves as if $k + 1$ threads were running concurrently, where k is a function of the asynchrony. In order to illustrate how the RC-simulation can be used, we describe a long-lived renaming implementation. By reducing the concurrency down to the asynchrony of the system, RC-simulation enables to obtain renaming quality that adapts linearly to the asynchrony.

5.3.7. Quantum Computing

In [1], we provide illustrative examples of distributed computing problems for which it is possible to design tight lower bounds for *quantum* algorithms without having to manipulate concepts from quantum mechanics, at all. As a case study, we address the following class of 2-player problems. Alice (resp., Bob) receives a boolean x (resp., y) as input, and must return a boolean a (resp., b) as output. A *game* between Alice and Bob is defined by a pair (δ, f) of boolean functions. The objective of Alice and Bob playing game (δ, f) is, for every pair (x, y) of inputs, to output values a and b, respectively, satisfying $\delta(a, b) = f(x, y)$, in *absence of any communication* between the two players, but in *presence of shared resources*. The ability of the two players to solve the game then depends on the type of resources they share. It is known that, for the so-called CHSH game, i.e., for the game $a \oplus b = x \wedge y$, the ability for the players to use entangled quantum bits (qubits) helps. We show that, apart from the CHSH game, quantum correlations do not help, in the sense that, for every game not equivalent to the CHSH game, there exists a classical protocol (using shared randomness) whose probability of success is at least as large as the one of any protocol using quantum resources. This result holds for both worst case and average case analysis. It is achieved by considering a model stronger than quantum correlations, the *non-signaling model*, which subsumes quantum mechanics, but is far easier to handle.

5.3.8. Distributed Decision and Verification

5.3.8.1. Randomization

In [12], we study the power of randomization in the context of locality by analyzing the ability to "boost" the success probability of deciding a distributed language. The main outcome of this analysis is that the distributed computing setting contrasts significantly with the sequential one as far as randomization is concerned. Indeed, we prove that in some cases, the ability to increase the success probability for deciding distributed languages is rather limited.

5.3.8.2. Model Variants

In a series of papers [14], [28], we analyze distributed decision in the context of various models for distributed computing.

In [28], we carry on the effort to bridging runtime verification with distributed computability, studying necessary conditions for monitoring failure prone asynchronous distributed systems. It has been recently proved that there are correctness properties that require a large number of opinions to be monitored, an opinion being of the form true, false, perhaps, probably true, probably no, etc. The main outcome of this paper is to show that this large number of opinions is not an artifact induced by the existence of artificial constructions. Instead, monitoring an important class of properties, requiring processes to produce at most k different values does require such a large number of opinions. Specifically, our main result is a proof that it is impossible to monitor k-set-agreement in an n-process system with fewer than $\min\{2k, n\} + 1$ opinions. We also provide an algorithm to monitor k-set-agreement with $\min\{2k, n\} + 1$ opinions, showing that the lower bound is tight.

Finally, in [14], we tackle *local distributed testing* of graph properties. This framework is well suited to contexts in which data dispersed among the nodes of a network can be collected by some central authority (like in, e.g., sensor networks). In local distributed testing, each node can provide the central authority with just a few information about what it perceives from its neighboring environment, and, based on the collected information, the central authority is aiming at deciding whether or not the network satisfies some property. We analyze in depth the prominent example of checking *cycle-freeness*, and establish tight bounds on the amount of information to be transferred by each node to the central authority for deciding cycle-freeness. In particular, we show that distributively testing cycle-freeness requires at least $\lceil \log d \rceil - 1$ bits of information per node in graphs with maximum degree d, even for connected graphs. Our proof is based on a novel version of the seminal result by Naor and Stockmeyer (1995) enabling to reduce the study of certain kinds of algorithms to order-invariant algorithms, and on an appropriate use of the known fact that every free group can be linearly ordered.

5.3.9. Voting Systems

In [44], [38], we consider a general framework for voting systems with arbitrary types of ballots such as orders of preference, grades, etc. We investigate their manipulability: in what states of the population may a coalition of electors, by casting an insincere ballot, secure a result that is better from their point of view?

We show that, for a large class of voting systems, a simple modification allows to reduce manipulability. This modification is *Condorcification*: when there is a Condorcet winner, designate her; otherwise, use the original rule.

When electors are independent, for any non-ordinal voting system (i.e. requiring information that is not included in the orders of preferences, for example grades), we prove that there exists an ordinal voting system whose manipulability rate is at most as high and which meets some other desirable properties. Furthermore, this result is also true when voters are not independent but the culture is *decomposable*, a weaker condition that we define.

Combining both results, we conclude that when searching for a voting system whose manipulability is minimal (in a large class of systems), one can restrict to voting systems that are ordinal and meet the Condorcet criterion.

In [35], we examine the geometrical properties of the space of expected utilities over a finite set of options, which is commonly used to model the preferences of an agent. We focus on the case where options are assumed to be symmetrical a priori, which is a classical neutrality assumption when studying voting systems. Specifically, we prove that the only Riemannian metric that respects the geometrical properties and the natural symmetries of the utility space is the round metric. Whereas Impartial Culture is widely used in Social Choice literature but limited to ordinal preference, our theoretical result allows to extend it canonically to cardinal preferences.

In [25], we study the manipulability of voting systems in a real-life experiment: electing the best paper in the conference Algotel 2012. Based on real ballots, we provide a quantitative study of the manipulability, as a function of the voting system used. We show that, even in a situation where all voting systems give the same winner by sincere voting, choosing the voting system is critical, because it has a huge impact on manipulability. In particular, one voting system fare way be better than the others: Instant-Runoff Voting.

5.4. Network Algorithms and Analysis

5.4.1. Bounds on the Cover Time in the Rotor-Router Model

In [23] and [33], we consider the *rotor-router mechanism*, which provides a deterministic alternative to the random walk in undirected graphs. In this model, a set of k identical walkers is deployed in parallel, starting from a chosen subset of nodes, and moving around the graph in synchronous steps. During the process, each node maintains a cyclic ordering of its outgoing arcs, and successively propagates walkers which visit it along its outgoing arcs in round-robin fashion, according to the fixed ordering. We consider the *cover time* of such a system, i.e., the number of steps after which each node has been visited by at least one walk, regardless of the starting locations of the walks. In the case of $k = 1$, Yanovski et al. (2003) and Bampas et al. (2009) showed that a single walk achieves a cover time of exactly $\Theta(mD)$ for any n-node graph with m edges and diameter D, and that the walker explores increasingly large Eulerian subgraphs before eventually stabilizes to a traversal of an Eulerian circuit on the set of all directed edges of the graph.

In [23], we provide tight bounds on the cover time of k parallel rotor walks in a graph. We show that this cover time is at most $\Theta(mD/\log k)$ and at least $\Theta(mD/k)$ for any graph, which corresponds to a speedup of between $\Theta(\log k)$ and $\Theta(k)$ with respect to the cover time of a single walk. Both of these extremal values of speedup are achieved for some graph classes. Our results hold for up to a polynomially large number of walks, $k = O(poly(n))$.

In [33], we perform a case study of cover time of the rotor-router, showing how the cover time depends on k for many important graph classes. We determine the precise asymptotic value of the rotor-router cover time for all values of k for degree-restricted expanders, random graphs, and constant-dimensional tori. For hypercubes, we also resolve the question precisely, except for values of k much larger than n. Our results can be compared to those obtained by Elsässer and Sauerwald (2009) in an analogous study of the cover time of k independent parallel random walks in a graph; for the rotor-router, we obtain tight bounds in a slightly broader spectrum of cases. Our proofs take advantage of a relation which we develop, linking the cover time of the rotor-router to the mixing time of the random walk and the local divergence of a discrete diffusion process on the considered graph.

5.4.2. Web Ranking and Aliveness

In [29] and [30], we investigate how to efficiently retrieve large portions of alive pages from an old crawl using orderings we called LiveRanks. Our work establishes the possibility of efficiently recovering a significant portion of the alive pages of an old snapshot and advocates for the use of an adaptive sample-based PageRank for obtaining an efficient LiveRank. Additionally, application field is not limited to Web graphs. It can be straightforwardly adapted to any online data with similar linkage enabling crawling, like P2P networks or online social networks.

5.4.3. Wireless Positioning

In [31], we consider how to construct a low-cost and efficient positioning system. We have proposed a new method called Two-Step Movement (2SM) to estimate the position of Mobile Terminal (MT). By exploiting useful information given by the position change of the device or user movement, this method can minimize the number of Reference Points (RP) required (*i.e.,* only one) in a localization system or navigation service and reduce system implementation cost. Analytical result shows that the user position can be derived, under noisy environment, with an estimation error about 10% of the distance between the RP and MT, or even less.

5.4.4. Content Centric Networking

Today's Internet usage is mostly centered around location-independent services. Because the Internet architecture is host-centric, content or service requests still have to be translated into locations, or the IP address of their hosts. This translation is realized through different technologies, e.g. DNS and HTTP redirection, which are currently implemented at the Application Layer. (ICN) proposes to evolve the current Internet infrastructure by extending the networking layer with name-based primitives.

In [45], we target the design and implementation of a content router, which is a network entity that implements *name-based forwarding*, or it can forward packets based on the content name they are addressed to. This work makes three major contributions. First, we propose an algorithm for name-based longest prefix match whose main novelty is the *prefix Bloom filter*, a Bloom filter variant that exploits the hierarchical nature of content prefixes. Second, a content router design that is compatible with both today's networking protocols and with widely used network equipments. Third, two innovative features that increase the scalability of a content router both in term of forwarding-information-base size and forwarding speed.

In the demonstration [34] held in the ICN conference, we demonstrate a high speed Information-Centric Network in a mobile backhaul setting. In particular, we emulate an information aware data plane and we highlight the significant benefits it provides in terms of both user experience and network provider cost in the backhaul setting. Our setup consists of high-speed ICN devices employed in a down-scaled realistic representation of a mobile backhaul topology, fed with traffic workloads characterized from Orange's mobile network. We compare numerical results activating and de-activating the ICN feature at run-time, showing the main differences between the two approaches. All the devices are implemented in a real high-speed multi-core equipment, and they are connected by means of internal port connections. Traffic is injected using a Traffic Generator which is implemented in the same architecture.

5.4.5. Information Dissemination

5.4.5.1. Dissemination with Noise or Limited Memory

In [26], we introduce the study of basic distributed computing problems in the context of noise in communication. We establish tight and almost tight bounds for the rumor spreading problem as well as for the majority-consensus problem.

In [11], we theoretically study a general model of information sharing within animal groups. We take an algorithmic perspective to identify efficient communication schemes that are, nevertheless, economic in terms of communication, memory and individual internal computation. We present a simple and natural algorithm in which each agent compresses all information it has gathered into a single parameter that represents its confidence in its behavior. Confidence is communicated between agents by means of active signaling. We motivate this model by novel and existing empirical evidences for confidence sharing in animal groups. We rigorously show that this algorithm competes extremely well with the best possible algorithm that operates without any computational constraints. We also show that this algorithm is minimal, in the sense that further reduction in communication may significantly reduce performances. Our proofs rely on the Cramér-Rao bound and on our definition of a Fisher Channel Capacity. We use these concepts to quantify information flows within the group which are then used to obtain lower bounds on collective performance.

5.4.5.2. Gossip and Rumor Spreading with Flooding

In [2], we address the flooding problem in dynamic graphs, where flooding is the basic mechanism in which every node becoming aware of an information at step t forwards this information to all its neighbors at all forthcoming steps $t' > t$. In particular, we show that a technique developed in a previous paper, for analyzing flooding in a Markovian sequence of Erdös-Rényi graphs, is robust enough to be used also in different contexts. We establish this by analyzing flooding in a sequence of graphs drawn independently at random according to a model of random graphs with given expected degree sequence. In the prominent case of power-law degree distributions, we prove that flooding takes almost surely $O(\log n)$ steps even if, almost surely, none of the graphs in the sequence is connected. In the general case of graphs with an arbitrary degree sequence, we prove several upper bounds on the flooding time, which depend on specific properties of the degree sequence.

5.4.6. Small-world Networks

In [9], we study decentralized routing in small-world networks that combine a wide variation in node degrees with a notion of spatial embedding. Specifically, we consider a variant of J. Kleinberg's grid-based small-world model in which (1) the number of long-range edges of each node is not fixed, but is drawn from a power-law probability distribution with exponent parameter $\alpha \geq 0$ and constant mean, and (2) the long-range edges are considered to be bidirectional for the purposes of routing. This model is motivated by empirical observations indicating that several real networks have degrees that follow a power-law distribution. The measured power-law exponent α for these networks is often in the range between 2 and 3. For the small-world model we consider, we show that when $2 < \alpha < 3$ the standard greedy routing algorithm, in which a node forwards the message to its neighbor that is closest to the target in the grid, finishes in an expected number of $O(\log^{\alpha-1} n \cdot \log\log n)$ steps, for any source–target pair. This is asymptotically smaller than the $O(\log^2 n)$ steps needed in Kleinberg's original model with the same average degree, and approaches $O(\log n)$ as α approaches 2. Further, we show that when $0 \leq \alpha < 2$ or $\alpha \geq 3$ the expected number of steps is $O(\log^2 n)$, while for $\alpha = 2$ it is $O(\log^{4/3} n)$. We complement these results with lower bounds that match the upper bounds within at most a $\log\log n$ factor.

5.4.7. Voting Systems and Path Selection in Networks

In [24], we apply our theoretical and experimental results on voting systems to a network use case: choosing a path in a network. In our model, nodes have an economical reward or cost for each possible path and they vote to elect the path. We show that the choice of the voting system has an important impact on the manipulability and the economical efficiency of this system. From both points of view, Instant-Runoff Voting gives the best results.

HIPERCOM2 Team

6. New Results

6.1. Highlights of the Year

- Hipercom 2 took part to the Inria-Industry meeting focusing on Telecommunications organized by Inria at Rocquencourt in November 2014. We presented a demonstration of the OCARI wireless sensor network.
- Hipercom 2 organized an Inria-DGA day "Software Defined Network (SDN) & MANET" at Paris in October 2014.

6.2. New Results about Wireless Sensor Networks

6.2.1. 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 2014, we study the issue of delay optimization and energy efficiency in grid wireless sensor networks (WSNs). We focus on STDMA (Spatial Reuse TDMA)) scheduling, where a predefined cycle is repeated, and where each node has fixed transmission opportunities during specific slots (defined by colors). We assume a STDMA algorithm that takes advantage of the regularity of grid topology to also provide a spatially periodic coloring ("tiling" of the same color pattern). In this setting, the key challenges are: 1) minimizing the average routing delay by ordering the slots in the cycle 2) being energy efficient. Our work follows two directions: first, the baseline performance is evaluated when nothing specific is done and the colors are randomly ordered in the STDMA cycle. Then, we propose a solution, ORCHID that deliberately constructs an efficient STDMA schedule. It proceeds in two steps. In the first step, ORCHID starts form a colored grid and builds a hierarchical routing based on these colors. In the second step, ORCHID builds a color ordering, by considering jointly both routing and scheduling so as to ensure that any node will reach a sink in a single STDMA cycle. We study the performance of these solutions by means of simulations and modeling. Results show the excellent performance of ORCHID in terms of delays and energy compared to a shortest path routing that uses the delay as a heuristic. We also present the adaptation of ORCHID to general networks under the SINR interference model.

6.2.2. 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 applications 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.

After a comprehensive survey on multichannel assignment protocols in wireless sensor networks, we study raw convergecast in multichannel wireless sensor networks (WSNs) where the sink may be equipped with multiple radio interfaces. We propose $Wave$,a simple, efficient and traffic-aware distributed joint channel and time slot assignment for raw convergecast. Our target is to minimize the data gathering delays and ensure that all packets transmitted in a cycle are delivered to the sink in this cycle, assuming no packet loss at the physical layer. We evaluate the number of slots needed to complete the convergecast by simulation and compare it to the optimal schedule and to a centralized solution. Simulations results indicate that our heuristic is not far from the optimal bound for raw convergecast. Unlike most previously published papers, $Wave$ does not suppose that all interfering links have been removed by channel allocation. In addition, $Wave$ is able to easily adapt to traffic changes. $Wave$ could be used to provide the schedule applied in the $802.15.4e$ TSCH based networks.

6.2.3. Optimized WSN Deployment

Participants: Ines Khoufi, Pascale Minet, Erwan Livolant.

This is a joint work with Telecom SudParis: Anis Laouiti.
We are witnessing the deployment of many wireless sensor networks in various application domains such as pollution detection in the environment, intruder detection at home, preventive maintenance in industrial process, monitoring of a temporary industrial worksite, damage assessment after a disaster.... Many of these applications require the full coverage of the area considered. With the full coverage of the area, any event occurring in this area is detected by at least one sensor node. In addition, the connectivity ensures that this event is reported to the sink in charge of analyzing the data gathered from the sensors and acting according to these data.

In the literature, many studies assume that this area is rectangular and adopt the classical deployment in triangular lattice that has been proved optimal. In real life, things are more complex. For instance, in an industrial worksite, the area to cover has an irregular shape with many edges and is not necessarily convex. Moreover, few papers take obstacles into account. Those that do assume that obstacles are constituted by a juxtaposition of rectangles that seems an unrealistic assumption. In real deployments, the shape of obstacles may be irregular. We distinguish two types of obstacles: the transparent ones like ponds in outdoor environment, or tables in an indoor site that only prevent the location of sensor nodes inside them; whereas the opaque obstacles like walls or trees prevent the sensing by causing the existence of hidden zones behind them: such zones may remain uncovered. Opaque obstacles are much more complex to handle than transparent ones and require the deployment of additional sensors to eliminate coverage holes. That is why we focus on the deployment of wireless sensor nodes in an arbitrary realistic area with an irregular shape, and with the presence of obstacles that may be opaque. Moreover, we propose a method that tends to minimize the number of sensor nodes needed to fully cover such an area.

Mobile robots can be used to deploy static wireless sensor nodes to achieve the coverage and connectivity requirements of the applications considered. Many solutions have been provided in the literature to compute the set of locations where the sensor nodes should be placed. We show how this set of locations can be used by a mobile robot to optimize its tour to deploy the sensor nodes to their right locations. In order to reduce both the energy consumed by the robot, its exposure time to a hostile environment, as well as the time at which the wireless network becomes operational, the optimal tour of the robot is this minimizing the delay. This delay must take into account not only the time needed by the robot to travel the tour distance but also the time spent in the rotations performed by the robot each time it changes its direction. This problem is called the Robot Deploying Sensor nodes problem, in short RDS. We first show how this problem differs from the well-known traveling salesman problem. We then propose an integer linear program formulation of the RDS problem. We propose various algorithms relevant to iterative improvement by exchanging tour edges, genetic approach and hybridization. The solutions provided by these algorithms are compared and their closeness to the optimal is evaluated in various configurations.

6.2.4. Sinks Deployment and Packet Scheduling for Wireless Sensor Networks

Participants: Nadjib Achir, Paul Muhlethaler.

The objective of this work is to propose an optimal deployment and distributed packet scheduling of multi-sink Wireless Sensors networks (WNSs). We start by computing the optimal deployment of sinks for a given maximum number of hops between nodes and sinks. We also propose an optimal distributed packet scheduling in order to estimate the minimum energy consumption. We consider the energy consumed due to reporting, forwarding and overhearing. In contrast to reporting and forwarding, the energy used in overhearing is difficult to estimate because it is dependent on the packet scheduling. In this case, we determine the lower-bound of overhearing, based on an optimal distributed packet scheduling formulation. We also propose another estimation of the lower-bound in order to simulate non interfering parallel transmissions which is more tractable in large networks. We note that overhearing largely predominates in energy consumption. A large part of the optimizations and computations carried out in this work are obtained using ILP formalization.

6.2.5. Security in wireless sensor networks

Participants: Selma Boumerdassi, Paul Muhlethaler.

Sensor networks are often used to collect data from the environment where they are located. These data can then be transmitted regularly to a special node called a *sink*, which can be fixed or mobile. For critical data (like military or medical data), it is important that sinks and simple sensors can mutually authenticate so as to avoid data to be collected and/or accessed by fake nodes. For some applications, the collection frequency can be very high. As a result, the authentication mechanism used between a node and a sink must be fast and efficient both in terms of calculation time and energy consumption. This is especially important for nodes which computing capabilities and battery lifetime are very low. Moreover, an extra effort has been done to develop alternative solutions to secure, authenticate, and ensure the confidentiality of sensors, and the distribution of keys in the sensor network. Specific researches have also been conducted for large-scale sensors. At present, we work on an exchange protocol between sensors and sinks based on low-cost shifts and xor operations.

6.2.6. Massive MIMO Cooperative Communications for Wireless Sensor Networks

Participants: Nadjib Achir, Paul Muhlethaler.

This work is a collaboration with Mérouane Debbah (Supelec, France).
The objective of this work is to propose a framework for massive MIMO cooperative communications for Wireless Sensor Networks. Our main objective is to analyze the performances of the deployment of a large number of sensors. This deployment should cope with a high demand for real time monitoring and should also take into account energy consumption. We have assumed a communication protocol with two phases: an initial training period followed by a second transmit period. The first period allows the sensors to estimate the channel state and the objective of the second period is to transmit the data sensed. We start analyzing the impact of the time devoted to each period. We study the throughput obtained with respect to the number of sensors when there is one sink. We also compute the optimal number of sinks with respect to the energy spent for different values of sensors. This work is a first step to establish a complete framework to study energy efficient Wireless Sensor Networks where the sensors collaborate to send information to a sink. Currently, we are exploring the multi-hop case.

6.2.7. 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.

When we use BMAC with opportunistic routing, one main advantage is that there is no transmission when there is no event detected in the network in contrast to RI-MAC where beacons of awaking nodes are periodically sent. However, when an event occurs in the area monitored, the end-to-end delay to deliver the alert packet to the sink is much greater with BMAC than with RI-MAC. This may pose problem to some real-time applications. We have propose a scheme where, instead of sending a long preamble to gather all the neighbor nodes, the packet is directly sent. The acknowledgement of the packet allows tthe sender to know whether (or not) the progression towards the destination is sufficient. If it is not the case the packet is sent again. More neighbor node will be awaken and the progression towards the destination will be improved. The selection of the relay terminates when the progression towards the destination is above a given threshold. Actually this relaying scheme encompasses two levels of opportunism. The first level consists in selecting only the awake nodes, the second level consists in selecting the best nodes among the awake nodes. We can show that doing so only slightly increase the number of hops to reach the sink whereas the delay per hop is largely reduced. Thus the end-to-end is very significantly reduced and we still have the property that there is no transmission when there is no event detected in the network.

6.3. Cognitive Radio Networks

6.3.1. 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.
The unlicensed spectrum bands become overcrowded causing an increased level of interference for current wireless sensor nodes. Cognitive Radio Sensor Networks (CRSNs) overcome this problem by allowing sensor nodes to access opportunistically the underutilized licensed spectrum bands. The sink assigns the spectrum holes to the secondary users (SUs). Therefore, it must rely on reliable information about the spectrum holes to protect the primary users (PUs). In 2013 we focused on the MultiChannel Time Slot Assignment problem (MC-TSA) in CRSN and proposed an Opportunistic centralized TIme slot assignment in COgnitive Radio sensor networks (OTICOR). This latter differs from the existing schemes in its ability to allow non-interfering cognitive sensors to access the same channel and time slot pair. OTICOR takes advantages of spatial reuse, multichannel communication and multiple radio interfaces of the sink. We proved through simulations that a smaller schedule length improves the throughput. Applying OTICOR, we show that, even in the presence of several PUs, the average throughput granted to SUs remains important. We also show how to get the best performances of OTICOR when the channel occupancy by PUs is known.

In 2014, we proposed two ways for the sink to determine the available channels and alert the SUs if an unexpected activity of PU occurs. Our objective is to design an algorithm able to detect the unexpected presence of PUs in the multi-hop network while maximizing the throughput. To achieve our goal, we propose an optimized version of our previous scheduling algorithm Opportunistic centralized TIme slot assignment in COgnitive Radio sensor networks (OTICOR). This algorithm takes advantage of the slots dedicated to the control period by allowing noninterfering cognitive sensors to access the control/data channel and time slot pair. We shown through simulations that using the control period for data transmission minimizes the schedule length and maximizes the throughput.

6.4. Mobile ad hoc and mesh networks

6.4.1. Development and implementation of a network coding module for NS3

Participants: Cédric Adjih, Ichrak Amdouni, Hana Baccouch.

DragonNet is a complete modular solution of network coding. This solution is responsible of coding, decoding, maintaining necessary information and the associated signaling. It is designed to be extensible. A variant of DragonNet was specified for wireless sensor networks and implemented.

As a follow-up to the ADT MOBSIM (and the previous module EyWifi), DragonNet was also integrated as a module for the NS-3 simulation tool.

6.4.2. Optimized Broadcast Scheme for Mobile Ad hoc Networks

Participants: Nadjib Achir, Paul Muhlethaler.

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. The obtained results demonstrate that our approach outperforms the classical flooding by increasing the delivery ratio and decreasing the number of required relays and thus the energy-cost.

6.5. Learning for an efficient and dynamic management of network resources and services

6.5.1. Learning in wireless sensor networks

Participants: Dana Marinca, Nesrine Ben Hassine, Pascale Minet, Selma Boumerdassi.

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.

In 2014, we applied on-line learning strategies to predict the quality of a wireless link in a WSN, based on the LQI metric and take advantage of wireless links with the best possible quality to improve the packet delivery rate. We model this problem as a forecaster prediction game based on the advice of several experts. The forecaster learns on-line how to adjust its prediction to better fit the environment metric values. A forecaster estimates the LQI value using the advice of experts. The model we propose learns on-line how to adapt to dynamic changes of the environment to compute efficient predictions. It presents a very good reactivity and adaptability. The simulations using traces collected in a real WSN based on the IEEE 802.15.4 standard have shown that the past time-windows which are effective for the prediction should have medium durations, about

200-400ms. The time windows durations less than 200ms do not give a good prediction, while durations larger than 400ms are efficient only in low variations environment. We note that these results strongly depend on the real traces, but the great advantage of the model is that it is self-adaptive to input traces profile. In this context, because of data normalization, the impact of loss functions is limited: entropy and square loss functions seem to give better and more stable predictions. Also, the experts prediction method should be adapted to traces profile. For low variation environment values, the average on past time windows is a good approximation. For high variation environment, a method predicting smoothed values close to minimum real values is more appropriate. Hence, the predicted values will be stabilized around the low values, avoiding estimations varying too much. Simulation results also show that for both types of experts (AMW and SES), the best expert depends on the phase considered. This is the reason why a forecaster is needed. Furthermore, the predictions of the EWA forecaster using SES experts are shown to be reactive and accurate. This combination minimizes the cumulated loss regarding the real LQI values, compared with any other combination such as EWA-AMW, BE-AMW and BE-SES, given by decreasing performance order.

6.5.2. Prediction and energy efficiency for datacenters

Participants: Dana Marinca, Nesrine Ben Hassine, Pascale Minet, Selma Boumerdassi.

The exponential development of Information and Communication Technologies (ICT) have led to an over consumption of services and data shared in networks. From computing in companies to unified communications through social networks and Internet of Things, the use of ICT a reach the highest level ever. The complexity involved by these different services reveals the limits of computing in companies and leads a majority of organisms to partially or completely host the management of there information system in data centers. The latter are larger and larger and are composed of buildings containing powerful computing equipments and air-conditionning systems. Data centers require a huge amount of energy. As an example, in 2014, the electric consumption of all date centers will be larger than 42 TWh, and after 2020 the CO2 production will be larger then 1.27 GTons, ie. more than the aeronautic industry (GeSI SMARTer 2020 report). These "frightening" figures led the research community to work on the management of energy consumption. Several tracks have been explored, among which the optimization of computation and load balancing of servers. At present, we work on tools dedicated to traffic prediction, thus allowing a better management of servers. Our work consists in modeling the traffic specific to data centers and apply different statistical prediction methods.

6.6. Vehicular Ad hoc NETworks (VANETs)

6.6.1. Congestion Control in VANETs

Participants: Paul Muhlethaler, Anis Laouiti.

We have reviewed the schemes of Congestion Control in VANETs for safety messages. The solutions proposed are: to adapt the generation rate, to adapt the transmission power or to adapt the carrier sense threshold. Some mechanisms employ different states depending on the channel load. Some other schemes use recursive adaptation of their parameters (e.g. LIMERIC). According to a few studies the recursive adaptation system provide a better adaptation of the VANET to the channel load. We will study how the transmission rate and the carrier sense threshold (or transmission power) can be best adapted in order to send CAM: Car Awareness Messages with the highest rate and to the furthest vehicles while maintaining the total load below a given threshold. We will also study the better combination of transmission rate and the carrier sense threshold for the CAM.

6.6.2. TDMA schemes for VANETs

Participants: Mohamed Hadded, Paul Muhlethaler, Anis Laouiti.

This is a joint work with Leila Saidane and Rachid Zabrouba from ENSI (Tunisia).

Vehicular Ad-hoc NETworks (VANETs) help improving traffic safety and efficiency. Each vehicle can exchange information to inform other vehicles about the current status of the traffic flow or a dangerous situation such as an accident. Road safety and traffic management applications require a reliable communication scheme with minimal transmission collisions, which thus increases the need for an efficient Medium Access Control (MAC) protocol. However, the MAC in a vehicular network is a challenging task due to the high speed of the nodes, frequent changes in topology, the lack of an infrastructure, and various QoS requirements. Recently several Time Division Multiple Access (TDMA)-based medium access control protocols have been proposed for vehicular ad hoc networks in an attempt to ensure that all the vehicles have enough time to send safety messages without collisions and reducing end-to-end delay and packet loss rate. We have identified the reasons for using the collision-free medium access control paradigm in VANETs. We have then presented a novel topology-based classification and we provide an overview of TDMA-based MAC protocols that have recently been proposed for VANETs. We have focus on the characteristics of these protocols as well as their benefits and limitations. Finally we have given a qualitative comparison, and we have discussed some open issues that need to be tackled in future studies to improve the performance of TDMA-based MAC protocols for vehicle to vehicle V2V communication.

MIMOVE Team

6. New Results

6.1. Introduction

MiMove's research activities in 2014 have focused on a set of areas directly related to the team's research topics. Hence, we have worked on Emergent Middleware (§ 6.3) and Service-oriented Computing in the Future Internet (§ 6.4), in relation to our research topic regarding Emergent Mobile Distributed Systems (§ 3.2). With respect to Large-scale Mobile Sensing & Actuation (§ 3.3), we have developed activities on Service-oriented Middleware for the Mobile Internet of Things (IoT) (§ 6.5), Composing Applications in the IoT (§ 6.6), and Lightweight Streaming Middleware for the IoT (§ 6.7). Last, our effort on Middleware for Mobile Social Networks (§ 6.8) is linked to our research on Mobile Social Crowd-sensing (§ 3.4).

Before presenting our new results in the areas mentioned above, we briefly discuss next the highlights of the year.

6.2. Highlights of the Year

This year has seen the following acknowledgments of the team's contributions:

- Valérie Issarny was distinguished as Chevalier de la Legion d'Honneur for her contributions to science and European scientific cooperation in research and education.
- One of the team's major publication by S. Ben Mokhtar, D. Preuveneers, N. Georgantas, V. Issarny, and Y. Berbers, titled "EASY: Efficient semAntic Service discoverY in pervasive computing environments with QoS and context support" [1], published in the Journal of Systems and Software (Volume 81, Issue 5), is one of the top ten (10) most cited papers among all the papers published by JSS in 2008.

6.3. Emergent Middleware

Participants: Emil Andriescu, Valérie Issarny, Thierry Martinez.

Our previous work on emergent middleware has focused on interconnecting 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 [18]. Our approach for the automated synthesis of mediators is performed through *interface matching*, which identifies the semantic correspondence between the actions required by one component and those provided by the other, followed by the *synthesis of correct-by-construction mediators*. To do so, we analyze the behaviors of components so as to generate the mediator that coordinates the matched actions in a way that guarantees that the two components progress and reach their final states without errors [2]. 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.

In our latest work, we have particularly focused on item (iii) above. We recognize that modern distributed systems and Systems of Systems (SoS) are built as a composition of existing components and services. As a result, systems communicate (either internally, locally or over networks) using protocol stacks of ever-increasing complexity whose messages need to be translated (i.e., interpreted, generated, analyzed and transformed) by third-party systems. We are particularly interested in the application of message translation to achieve protocol interoperability via protocol mediators. We observe that current approaches are unable to provide an efficient solution towards reusing message translators associated with the message formats composed in protocol stacks. Instead, developers must write ad hoc "glue-code" whenever composing two or more message translators.

Ideally, message translators may be developed by separate parties, using various technologies, while developers should be able to compose them using an easy to use mechanism. However, parsers are monolithic and tightly constructed, which often makes it impossible to combine them, knowing that combining two unambiguous grammars (corresponding to two arbitrary parsers) may result in an ambiguous grammar, and that the ambiguity detection problem for context-free grammars is undecidable in the general case.

In addition to parser composition, the data structures of the parsing output must be manually defined, integrated and harmonized with the target systems (i.e., in this case, the Mediation Engine). As far as we know, the problem of inferring the output schema (or the data type) of an arbitrary tree transformation has not yet been solved, while it is known that, in general, a transformation might not be recognizable by a schema.

Following the challenges above, in [17], we make two major contributions to the issue of systematic message translation for modern distributed systems:

1. Starting from the premise that "off-the-shelf" message translators for individual protocols are readily available in at least an executable form, we propose a solution for the automated composition of message translators. The solution simply requires the specification of a composition rule that is expressed using a subset of the navigational core of the W3C XML query language XPath.
2. We provide a formal mechanism, using tree automata, which based on the aforementioned composition rule, generates an associated AST *data-schema* for the translator composition. This contribution enables the inference of correct data-schemas, relieving developers from the time-consuming task of defining them. On a more general note, the provided method solves the type inference problem for the *substitution* class of tree compositions in linear time on the size of the output. The provided inference algorithm can thus be adapted to a number of applications beyond the scope of this work, such as XML Schema inference for XSLT transformations.

The composition approach that we introduced functions as a purely "black-box" mechanism, thus allowing the use of third-party parsers and message serializers independently of the parsing algorithm they use internally, or the method by which they were implemented/generated. Our solution goes beyond the problem of translator composition by inferring AST data-schemas relative to translator compositions. This feature allows newly generated translators to be seamlessly (or even automatically) integrated with existing systems, and most notably our protocol mediation engine [2].

6.4. Service-oriented Computing in the Future Internet

Participants: Georgios Bouloukakis, Nikolaos Georgantas, Valérie Issarny, Ajay Kattepur, Raphael de Aquino Gomes, Rachit Agarwal.

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 diverse systems as application components that interact via standard middleware protocols. However, the heterogeneous nature of such systems leads to choreographies that do not only include conventional services, but also sensor-actuator networks, databases and service feeds. We reason about the behavior of such systems by introducing abstract middleware connectors that follow base interaction paradigms, such as client-service (CS), publish-subscribe (PS) and tuple space (TS). These heterogeneous connectors are made interoperable through a service bus connector, the *eXtensible Service Bus* (XSB) [11].

In previous work, we identified and verified the behavioral semantics of the XSB connector derived from the interconnection of base connectors, and introduced a method for constructing protocol converters enabling this interconnection. We implemented our XSB solution into an extensible development and execution platform for application and middleware designers. We also provided a lightweight implementation of the XSB, the *Light Service Bus* (LSB), appropriate for resource-constrained environments and systems. Next, leveraging on the functional interoperability across interaction paradigms offered by the XSB, we initiated our study of end-to-end Quality of Service (QoS) properties of choreographies, where in particular we focus on the effect of middleware interactions on QoS.

Building on the above results, we refine our analysis of QoS on top of the identified interaction paradigms. We have introduced a motivating application scenario inspired from the *2014 D4D Challenge*[0]. More specifically, *Data for Development Senegal* is an innovation challenge on ICT Big Data for the purposes of societal development. Mobile network provider Sonatel (part of the Orange Group) has made anonymous data extracted from the mobile network in Senegal available to international research laboratories, encouraging research related to the development and welfare of the local population.

Our scenario targets the development of an application platform for citywide and countrywide transport information management relying on mobile social crowd-sensing. This takes into account the particular context and constraints in Senegal. More specifically, the local transportation system, although developing, still consists of many unplanned and informal settlements with unreliable services and infrastructure. Additionally, despite wide use of mobile phones in the country, mobile Internet access remains limited, making SMS the only alternative for data access for a large part of the population. Our proposition aims to complement the scarce authoritative transport information coming from structured information sources and compensate for the lack of such information. In particular, in our approach we intend to study and experiment with appropriate interaction paradigms (CS, PS, TS) on top of 3G/2G/SMS data connections, further depending on the specific application and data. We are especially interested in interaction adaptation depending on the network conditions (e.g., switching to SMS-based protocol when the 3G/2G network is unavailable).

We have taken a first step towards enabling such an application platform. This consists in evaluating the publish/subscribe interaction style in a large-scale setting where resources of mobile users are limited, which translates into limited and intermittent connectivity in the system. Additionally, such an application platform must guarantee that the sensing data is processed and delivered to the corresponding mobile users *on-time*, despite the intermittent connectivity of the latter. We have opted for the publish/subscribe paradigm, as it is deemed appropriate for loose spatio-temporal interaction between mobile entities.

In particular, we introduce a queueing network model for the end-to-end interaction within a large-scale mobile publish/subscribe system. We leverage the *D4D dataset* provided by Orange Labs to parametrize this model. We then develop a simulator named *MobileJINQS*[0] that implements our model and uses the dataset traces as realistic input load to the system model over the time span of a whole year. Prior to this, we extensively analyze the D4D dataset in order to identify the data that we are interested in and infer primary results [0]. Based on the results of our simulation-based experiments, we thoroughly evaluate the behavior of the publish/subscribe system and identify ways of tuning the system parameters in order to satisfy certain design requirements. More precisely, we provide results of simulations of our publish/subscribe system with varied incoming loads, service delays and event lifetime periods. We use connection data of various pairs of mobile network antennas to derive realistic traces for both incoming loads and service delays. System or application designers are able to tune the system by selecting appropriate lifetime periods. We demonstrate that varying incoming loads and service delays have a significant effect on response time. By properly setting event lifetime spans, designers can best deal with the tradeoff between freshness of information and information delivery success rates. Still, both of these properties are highly dependent on the dynamic correlation of the event input flow and delivery flow processes, which are intrinsically decoupled.

[0] http://www.d4d.orange.com/en/home
[0] http://xsb.inria.fr/d4d#mobilejinqs
[0] http://xsb.inria.fr/d4d

Our future work includes comparison of the publish/subscribe interaction paradigm with other interaction paradigms (client-server, tuple space), in relation with the network access capacity and the application requirements. Also, we intend to study the response time and success rate for the various combinations of antennas in more fine-grained scales (e.g., check what their evolution is over one day).

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

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

The Internet of Things (IoT) is characterized by a wide penetration in the regular user's life through 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 which lead to high communication and computation costs especially with the anticipated large volumes of data to exchange, ii) providing interoperability across the heterogeneous Things which host sensors and actuators providing services and producing data that follow different format/schema specifications, 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. Our work on the revisited Thing-based SOA is detailed in [9], [23], [15]. 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 protocols to provide, not all, but only a sufficient subset of services that can best approximate the result that is being sought after [23], [15] based on a predefined set of requirements such us sensing coverage of the area of interest and the location of the Things. By limiting the participation of Things, the communication costs and volumes of data to process are decreased without jeopardising the quality of the outcome.

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). To that end, we model our composition specification as mathematical formulas defined semantically within a dedicated ontology. 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 and relying on semantic technologies to identify the most appropriate services to compose.

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. It revisits Service-Oriented access and leverage semantic technologies 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 real-time query-based access to remote services and to locally hosted services.

To assess the validity of our proposed architecture, we provide a prototype implementation of MobIoT(§ 5.4) along with a set of extensive evaluations that demonstrate, not only the feasibility of our approach, but also the resulting quality of the discovery approach, along with its scalability, as compared to a regular SOA-based approach.

6.6. Composing Applications in the Internet of Things

Participants: Aness Bajia, Animesh Pathak, Françoise Sailhan.

Resilient computing is defined as the ability of a system to remain dependable when facing changes. To mitigate faults at runtime, dependable systems embark fault tolerance mechanisms such as replication techniques. These mechanisms have to be systematically and rigorously applied in order to guarantee the conformance between the application runtime behavior and its dependability requirements.

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 application-level fault tolerance requirements, as well as by developers of fault tolerance protocols to identify the abilities and requirements of their techniques; ii) a runtime system that employs adaptive fault tolerance (AFT) to provide fault tolerance to the networking sensing application; and iii) compilation techniques to instantiate and map tasks as needed to satisfy the requirements of the application for a given deployment. Through our work [26], we demonstrate that our approach provides this much-needed feature to networked sensing applications with negligible development- and minimal performance- overhead.

Complementary to the above, we have proposed task mapping algorithms to satisfy those requirements through a constraint programming approach [24]. 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 been continually incorporating our research results in the above areas into *Srijan* (§ 5.5), which provides an easy-to-use graphical front-end to the various steps involved in developing an application using the ATaG macroprogramming framework.

6.7. Lightweight Streaming Middleware for the Internet of Things

Participants: Benjamin Billet, Valérie Issarny.

The IoT raises many challenges related to its very large scale and high dynamicity, as well as the great heterogeneity of the data and systems involved (e.g., powerful versus resource-constrained devices, mobile versus fixed devices, continuously-powered versus battery-powered devices, etc.). These challenges require new systems and techniques for developing applications that are able to: (i) collect data from the numerous data sources of the IoT, and (ii) interact both with the environment using the actuators and with the users using dedicated GUIs. Given the huge volume of data continuously being produced by sensors (measurements and events), we must consider: (i) data streams as the reference data model for the IoT and, (ii) continuous processing as the reference computation model for processing these data streams. Moreover, knowing that privacy preservation and energy consumption are increasingly critical concerns, we claim that all the Things should be autonomous and work together in restricted areas as close as possible to the users rather than systematically shifting the computation logic into powerful servers or into the cloud.

Toward that goal, we have been developing Dioptase [3], a service-oriented middleware for the IoT, which aims to integrate the Things and their streams into today's Web by presenting sensors and actuators as Web services. The research work around the Dioptase middleware consists in designing new service-oriented architectures where services continuously process data streams instead of finite datasets. In this context, new composition mechanisms are investigated in order to provide a way to describe complex fully-distributed stream-based tasks 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 communicate using data streams. Dynamic tasks are then described in a lightweight DSL, called *DiSPL*, which is directly interpreted by the middleware and provides specific primitives to manipulate data streams.

As part of the design of such composition mechanisms, we have been investigating the problem of task mapping and automated deployment, which basically consists of mapping a set of tasks onto a set of nodes. Given the specific challenges introduced by the IoT, we worked on a new 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). This formalization, called *Task Graph to Concrete Actions (TGCA)* [19], 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.

6.8. Middleware for Mobile Social Networks

Participants: Animesh Pathak, George Rosca.

As recent trends show, online social networks (OSNs) are increasingly turning mobile and further calling for decentralized social data management. This trend is only going to increase in the near future, based on the increased activity, both by established players like Facebook and new players in the domain such as Google, Instagram, and Pinterest. Modern smart phones can thus be regarded as *social sensors*, collecting data not only passively using, e.g., Bluetooth neighborhoods, but actively in the form of, e.g., "check-in"s by users to locations. The resulting (mobile) social ecosystems are thus an emergent area of interest.

The recent years have seen three major trends in the world of online social networks: *i)* users have begun to care more about the privacy of their data stored by large OSNs such as Facebook, and have won the right (at least in the EU) to remove it completely from the OSN if they want to; *ii)* OSNs are making their presence felt beyond casual, personal interactions to corporate, professional ones as well, starting with LinkedIn, and most recently with the purchase by Microsoft of Yammer, the enterprise social networking startup, and the launch of Google Plus for enterprise customers; and *iii)* users are increasingly using the capabilities of their (multiple) mobile devices to enrich their social interactions, ranging from posting cellphone-camera photos on Instagram to "checking-in" to a GPS location using Foursquare.

In view of the above, we envision that in the near future, the use of ICT to enrich our social interactions will grow (including both personal and professional interactions), both in terms of size and complexity. However current OSNs act mostly like data silos, storing and analyzing their users' data, while locking in these users to their servers, with non-existent support for federation; this is reminiscent of the early days of email, where one could only email those who had accounts on the same Unix machine. The knee-jerk reaction to this has been to explore completely decentralized social networks, which give the user complete control over and responsibility of their social data, while resorting to peer-to-peer communication protocols to navigate their social networks. Unfortunately, there are few techniques available to reconcile with the fact that the same user might have multiple devices, or that it is extremely resource-consuming to perform complex analysis of social graphs on small mobile devices.

Our view lies somewhere in the middle of the two extremes, taking inspiration from the manner in which users currently use email. While their inboxes contain an immense amount of extremely personal data, most users are happy to entrust it to corporate or personal email providers (or store and manage it individually on their personal email servers) all the while being able to communicate with users on any other email server. The notion of *Federated Social Networks* (FSNs) —already gaining some traction— envisions a similar ecosystem where users are free to choose OSN providers which will provide storage and management of their social information, while allowing customers using different OSN providers to interact socially. Such a federation can be beneficial in three major ways, among others: *i)* it allows users to enjoy properties such as reliability,

availability, and computational power of the hosting infrastructure of their choice, while not being locked down in terms of whom they can communicate with; *ii)* much like spam filtering services provided by modern email providers, that are tuned by feedback from their users, FSN users can benefit from the behavior of others sharing the same OSN provider [0]; and *iii)* this fits perfectly with enterprise needs, where ad-hoc teams can be formed across corporate OSN providers of two organizations to work on a joint project.

In [30], we presented a set of requirements, followed by a survey of the state of the art in social networking solutions, with a special focus on their ability to support rich privacy and access control policies in federated settings. Through this extensive analysis we offer a broad vision on existing social networking platforms, protocols involved but also their privacy and access policies. By doing so, we identify the main components of a federated social platform together with presenting the current trends in standards and security paradigms underlying actual open source solutions which offers their implementation, and finally provides recommendations on constructing such systems. Our research is continually being incorporated into the Yarta middleware for mobile social networking(§ 5.7).

[0]This also gives an incentive to commercial OSN providers to provide value-added services.

MUSE Team

6. New Results

6.1. Pinpointing Home and Access Network Delays Using WiFi Neighbors

Participants: Lucas Di Cioccio (LIP6/Technicolor), Martin May (Technicolor), Jim Kurose (University of Massachusetts, Amherst), Renata Teixeira

Home Internet users and Internet access providers need tools to assist them in diagnosing and troubleshooting network performance problems. Today, expert users may rely on simple techniques using round-trip-measurements to local and remote points to locate delays on an end-to-end path. Unfortunately, round-trip measurements do not provide accurate diagnoses in the presence of asymmetric link capacities and performance, which is often the case in residential access. Our work [8] introduces *neighbor-assisted delay diagnosis* (NADD) - an approach for pinpointing the location of delays (among the home, access, and wide-area network), leveraging end-host multi-homing capabilities. NADD runs on an end host connected simultaneously to the home gateway and to a neighbor WiFi access point. Our evaluation shows that NADD efficiently detect and distinguish uplink and downlink delays with small error. In addition, we learn from a proof-of-concept deployment in five homes in France that our techniques can work "in the wild." Technicolor filed a patent on this work [8].

6.2. Locating Throughput Bottlenecks in Home Networks

Participants: Srikanth Sundaresan (ICSI), Nick Feamster (Princeton), Renata Teixeira

We developed *WTF (Where's The Fault?)* [4], a system that localizes performance problems in home and access networks. We implement WTF as custom firmware that runs in an off-the-shelf home router. WTF uses timing and buffering information from passively monitored traffic at home routers to detect both access link and wireless network bottlenecks. The Federal Communication Commission (FCC) in the United States deployed WTF in 3000 homes for a few days in November 2014. We are currently analyzing the resulting dataset to help shed light on common pathologies that occur in home networks.

6.3. Measuring the Performance of User Traffic in Home Wireless Networks

Participants: Srikanth Sundaresan (ICSI), Nick Feamster (Princeton), Renata Teixeira

This work [5] studies how home wireless performance characteristics affect the performance of user traffic in real homes. Previous studies have focused either on wireless metrics exclusively, without connection to the performance of user traffic; or on the performance of the home network at higher layers. In contrast, we deploy a passive measurement tool on commodity access points to correlate wireless performance metrics with TCP performance of user traffic. We implement our measurement tool, deploy it on commodity routers in 66 homes for one month, and study the relationship between wireless metrics and TCP performance of user traffic. We find that, most of the time, TCP flows from devices in the home achieve only a small fraction of available access link throughput; as the throughput of user traffic approaches the access link throughput, the characteristics of the home wireless network more directly affect performance. We also find that the 5 GHz band offers users better performance better than the 2.4 GHz band, and although the performance of devices varies within the same home, many homes do not have multiple devices sending high traffic volumes, implying that certain types of wireless contention may be uncommon in practice.

6.4. Characterizing Bufferbloat and its Impact at End-hosts

Participants: Stephane Wustner, Jaideep Chandrashekar (Technicolor), Renata Teixeira

While, on routers and gateways, buffers on forwarding devices are required to handle bursty Internet traffic, overly large or badly sized buffers can interact with TCP in undesirable ways. This phenomenon is well understood and is often called "bufferbloat". Although a number of previous studies have shown that buffering (particularly, in home) can delay packets by as much as a few seconds in the worst case, there is less empirical evidence of tangible impacts on end-users. In [3], we develop a modified algorithm that can detect bufferbloat at individual end-hosts based on passive observations of traffic. We then apply this algorithm on packet traces collected at 55 end-hosts, and across different network environments. Our results show that 45 out of the 55 users we study experience bufferbloat at least once, 40% of these users experience bufferbloat more than once per hour. In 90% of cases, buffering more than doubles RTTs, but RTTs during bufferbloat are rarely over one second. We also show that web and interactive applications, which are particularly sensitive to delay, are the applications most often affected by bufferbloat.

6.5. Measuring and Characterising User Online Activity

Participants: Omayma Belkadi, Mauricio Santoro, Anna-Kaisa Pietilainen, Renata Teixeira

The goal of our work is to identify what people are doing online (or the online user activity) from passively collected network traffic traces. Our analysis of network traffic and application information from 12 end-hosts shows that this task is challenging because there are often many applications running on each user's device, whereas the user is only interacting with one application at a time. Our work with two master students presents the first evaluation of the set of features computable from network traffic alone that can help distinguish user activity traffic from all other traffic flows [6], [10]. We obtain ground truth on user activities and network traffic traces in a controlled setting, and complement this dataset with traces collected by the HostView monitoring tool on the devices of 12 users over several months. We develop simple heuristics to extract user activities for the HostView dataset based on the foreground application and on keyboard/mouse activity. Then, we analyze which network traffic features allow us to distinguish between online user activity and background network traffic. Features related to traffic volumes and timings show the most significant differences.

6.6. WeBrowse: a Passive Content Curation System Based on HTTP Logs

Participants: Giuseppe Scavo, Zied Ben Houidi (Alcatel-Lucent), Renata Teixeira, Stefano Traverso (Politecnico di Torino), Marco Mellia (Politecnico di Torino)

Content curation refers to the act of assisting users to identify relevant and interesting information in the overwhelming amount of online content available today. Existing curation services rely either on experts or on crowdsourcing to promote content. This work designs, implements, and evaluates WeBrowse, the first passive crowdsourced content curation system. WeBrowse requires no active user engagement to promote content. Instead, it extracts the URLs users visit from traffic traversing an ISP network to identify popular and interesting content. A key challenge to design such a passive curation system is to process network traffic in real-time to identify the small set of URLs that are interesting to users. WeBrowse contains a set of heuristics to identify the set of URLs users visit and to select the subset that are interesting, while preserving their privacy at the same time. We prototype WeBrowse and evaluate it using traces collected at a large European ISP, and in a deployment in a large campus network. We have tested and improved WeBrowse with a small number of users from September 2014 to January 2015. The plan is to announce WeBrowse to all users of the campus network early 2015 to get feedback on their experience with the system.

Available at: http://tstat.polito.it/netcurator/

RAP Project-Team

4. New Results

4.1. Random Graphs

Participants: Nicolas Broutin, Henning Sulzbach.

4.1.1. Universality of scaling limits of random graphs

Random graphs are one of the most studied models of networks, and they turn out to be related to crucial questions in physics about the behaviour of matter at the phase transition, or in combinatorial optimization about the hardness of computation. In recent years, we have constructed the scaling limit of the classical Erdos-Renyi random graph model, and conjectured that this limit also happened to be universal.

The funding of the Associated Team RNA has permitted to invite Shankar Bhamidi. During his visit, we have worked and found a new way to construct the scaling limit of random graph processes in the critical window. This method is especially important since it is robust enough to prove universality of the limit, that is that many models have the same limit. The method relies on the dynamics of the coalescence of clusters as the edges are added, and allows us to hope for proofs that would be able to treat the more complex geometric models.

4.1.2. Cutting down random tree and the genealogy of fragmentations

The study of the internal structure of random combinatorial object such as graphs and trees led to question about whether such objects exhibit invariance by certain complex surgical operations (disconnect some pieces, and re-attach them somewhere else). In the context of graphs, this is related to the so-called self-organized criticality: certain distributions that yield fractal objects should naturally appear in nature because they are the fixed points of some recombination procedures. In the context of trees, it turns out that certain fragmentations arising when chopping off a random tree have a genealogy that has the same distribution as the original tree. We have investigated this with Minmin Wang, and obtained results about p-trees and the genealogy of the fragmentation on Aldous' celebrated continuum random tree. These may also be interpreted in terms of complex path transformations for Brownian excursions and other random processes with exchangeable increments, and hence relate to very classical questions in probability theory.

4.1.3. New encodings for combinatorial coalescent processes

In 2013, we had constructed the scaling limit of the minimum spanning tree of a complete graph using crucial information about the scaling limit of random graphs, and especially about the way the cluster merge as the edges are added in the graph. With J.-F. Marckert (LaBRI, Bordeaux) we have found a novel construction of the important multiplicative coalescent that describes how the connected components of a random graph coalesce as the edges are added. This unveils yet more interesting links between the minimum spanning tree and the random graph, since Prim's celebrated algorithm is used to construct a consistent ordering of the vertices that ensures that the connected components are intervals.

4.1.4. Navigation in random Delaunay triangulations

Navigation or routing algorithms are fundamental routines: in order to solve many problems, one of the first steps consists in locating a node in a data structure. Unfortunately, the current algorithms are based on heuristics and very few rigorous results about the performance of such algorithms are known when the model for data is more realistic than the worst-case.

With O. Devillers and R. Hemsley, we have initiated a program that aims at finding rigourous estimates for the performance of routing algorithms in geometric structures such as Delaunay tesselations. So far we have managed to develop some tools that permitted us to analyse a simple algorithm. Although this algorithm has been designed for most of the analysis to work, this work paves the way towards the rigorous analysis of other more natural and widely used algorithms.

4.1.5. Connectivity and sparsification of sparse wireless networks

Many models of wireless networks happen to be connected only when the average degree is tending to infinity with the size of the network, more precisely when it is about the logarithm of the number of nodes. This raises questions about the potential issues in scaling such models. With L. Devroye (McGill) and G. Lugosi (ICREA and Pompeu Fabra), we have worked at analysing models in which we try to construct connected or almost connected networks in a distributed way (that is that no global optimization is allowed in designing the network, and every device should proceed in the same way to choose its neighbors). We have managed to analyse an algorithm for constructing such a network, and to obtain tight results about the number of links that a typical device should have in order for the global network to be connected. We further proved that this is asymptotically optimal when one only requires that most nodes should be in the same connected component.

4.2. Resource Allocation Algorithms in Large Distributed Systems

Participants: Christine Fricker, Philippe Robert, Guilherme Thompson.

This is a collaboration with Fabrice Guillemin from Orange Labs which started in February 2014.

4.2.1. Controlling impatience in cellular networks using QoE-aware radio resource allocation

Impatience of users when using a data service has a major impact on the quality of service offered by telecommunication networks, especially in cellular networks with scarce radio resources. Impatience is negative for users, it is due to many factors related to the performance of servers, customer devices, etc., but also to bandwidth sharing in the network.

While impatience can be seen as a negative phenomenon, it can also be used as a lever to discourage customers when the system becomes too much overloaded. This can be achieved in cellular networks by modulating the capacity available to customers being at a certain distance of the antenna. This general idea can be applied in several manners and can be viewed as a network optimization mechanism. In this paper, we reuse the general framework of α-fair scheduler in order to perform this control. This has the advantage of being easy to implement in realistic settings as α-fair schedulers (and especially the Proportional Fair (PF) one) are widely adopted in mobile networks. This also reduces the dimension of our problem as it narrows the optimization problem to the tuning of a single parameter α.

In order to achieve this goal, we first derive a model for reneging probabilities under a general α-fair scheduler. In particular, we consider a heavy load regime and develop a fluid flow analysis of impatience in cellular networks. We notably establish a fixed point formulation for the computation of the reneging probability and introduce a new metric, namely QoE perturbation, expressing how much a particular flow impacts the reneging probability in the system. We then use this QoE perturbation metric to design of a new radio resource management scheme that controls the parameter of the scheduler in order to reduce the global reneging in the system. For instance, recognizing that customers far from the base station degrade the global performance of the system, impatience and α-fair scheduling can be used to discourage those customers and in some sense to perform an implicit admission control in order to optimize the use of radio resources.

4.2.2. Resource Allocation in Large Data Centers

The goal of this study is to investigate the design of allocation algorithms of requests requiring different classes of quality of video streams as well as their performances. The class of algorithms considered may downgrade the quality of some of the transmission to maximize the utilization of the servers.

4.3. Stochastic networks: large bike sharing systems

Participants: Christine Fricker, Hanène Mohamed, Cédric Bourdais, Yousra Chabchoub.

Vehicle sharing systems are becoming an urban mode of transportation, and launched in many cities, as Velib' and Autolib' in Paris. One of the major issues is the availability of the resources: vehicles or free slots to return them. These systems became a hot topic in Operation Research and now the impact of stochasticity on the 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 stochastic model is the first studying the impact of the finite number of spots at the stations on the system behavior.

With Danielle Tibi, we use limit local theorems to obtain the asymptotic stationary joint distributions of several node (station or route) states when the system is large (both numbers of stations and bikes), also in the case of finite capacities of the stations. This gives an asymptotic independence preperty for node states. This widely extends the existing results on heterogeneous bike-sharing systems.

Second we investigate the impact of finite capacity of stations and reservation in car-sharing systems. The large-scale asymptotic joint stationary distribution of the numbers of vehicles and reserved parking places is given as the joint distribution in a tandem of queues with a constrained total capacity where rates are solutions of a system of two fixed point equations. Analytical expressions are given for performance in light and heavy traffic cases. As expected, reservation impact drastically increases with traffic. Even if the equilibrium is identified and analyzed, the question of convergence is still open.

JC Decaux provides us data describing Velib' user trips. These data are useful to measure the system behavior. With Yousra Chabchoub, we test clustering to obtain a typology of the stations. Then we focus on the resources availability (free docks and available bikes) and separate the Velib' stations into three clusters (balanced, overloaded and underloaded stations), using Kmeans clustering algorithm, along with the Dynamic Time Wraping (DTW) metric. We choose to update the centers of the clusters using the efficient Dtw Barycenter Averaging (DBA) method.

4.4. Scaling Methods

Participants: Philippe Robert, Wen Sun, Mohammadreza Aghajani.

4.4.1. Fluid Limits in Wireless Networks

This is a collaboration with Amandine Veber (CMAP, École Polytechnique). The goal is to investigate the stability properties of wireless networks when the bandwidth allocated to a node is proportional to a function of its backlog: if a node of this network has x requests to transmit, then it receives a fraction of the capacity proportional to $\log(1+x)$, the logarithm of its current load. 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.

4.4.2. The Time Scales of a Transient Network

The Distributed Hash Table (DHTs) consists of a large set of nodes connected through the Internet. Each file contained in the DHT is stored in a small subset of these nodes. Each node breaks down periodically and it is necessary to have back-up mechanisms in order to avoid data loss. A trade-off is necessary between the bandwidth and the memory used for this back-up mechanism and the data loss rate. Back-up mechanisms already exist and have been studied thanks to simulation. To our knowledge, no theoretical study exists on this topic. With a very simple centralized model, we have been able to emphasise a trade-off between capacity and life-time with respect to the duplication rate. From a mathematical point of view, we are currently studying different time scales of the system with an averaging phenomenon.

4.5. Stochastic Models of Biological Networks

Participants: Renaud Dessalles, Sarah Eugene, Emanuele Leoncini, Philippe Robert.

4.5.1. Stochastic Modelling of self-regulation in the protein production system of bacteria

This is a collaboration with Vincent Fromion from INRA Jouy-en-Josas, which started on December 2014.

In procaryots cells (e.g. E. Coli. or B. Subtilis) the protein production system has to produce in a cell cycle (i.e. less than one hour) more than 10^6 molecules of more than 2500 kinds, each having different level of expression. The bacteria uses more than 85% of its resources to the protein production. Gene expression is a highly stochastic process: bacteria sharing the same genome, in a same environment will not produce exactly the same amount of a given protein. Some of this stochasticity can be due to the system of production itself: molecules that take part in the production process move freely into the cytoplasm and therefore reach any target in the cell after some random time; some of them are present in so much limited amount that none of them can be available for a certain time; the gene can be deactivated by repressors for a certain time etc...

We study the integration of several mechanisms of regulation and their performances in terms of variance and distribution. All molecules are supposed to move freely into the cytoplasm, it is assumed that the the encounter time between a given entity and its target is exponentially distributed.

4.5.1.1. Transcription-translation model for all proteins

The first model that has been studied integrates the production of all the proteins. Each gene has to be transcribed in mRNA and each mRNA has to be translated in protein. The transcription step needs a RNA-Polymerase molecule that is sequestered during the time of elongation. Likewise, each mRNA needs a ribosome in order to produce a protein. RNA-Polymerases/Ribosomes are present in limited amount and the genes/mRNAs sequester these molecules during the whole the time of elongation. Finally each mRNA has an exponentially distributed lifetime with an average value of 4 min and the proteins disappear at a rate of one hour, hence simulating the global dilution in the growing bacteria.

This global sharing of Ribosomes/RNA-Polymerases among all proteins induces a general regulation: each gene competing to each other to have access to these common resources. Because of the parameters of affinity (between gene and RNA-Poymerase and between mRNA and ribosome) are specific to each gene, it allows a large range of average protein production but induce some noise, especially for highly expressed proteins.

We developed a Python simulation, and using the biological experiments of Tanichuchi et al. (2010), and we have investigated a biologically coherent range of parameters. By making the simulations, we have been able to reproduce certain aspects of the biological measures, especially for the high amount of noise for well expressed proteins.

4.5.1.2. Simple feedback model

We have also investigated the production of a single protein, with the transcription and the translation steps, but we also introduced a direct feedback on it: the protein tends to bind on the promoter of its own gene, blocking therefore the transcription. The protein remains on it during an exponential time until its detachment caused by thermal agitation.

The mathematical analysis aims at understanding the nature of the internal noise of the system and to quantify it. We try to determine if, for instance, for the same average protein level, the feedback permits a noise reduction of protein distribution compared to the "open loop" model; or if it rather allows a better efficiency in case of a change of command for a new level of production (due, for example, to a radical change in the environment) by reducing the respond time to reach this new average.

4.5.2. Stochastic Modelling of Protein Polymerization

This is a collaboration with Marie Doumic, Inria MAMBA team.

Our work focuses on the study of the polymerization of protein. This phenomenon is involved in many neurodegenerative diseases such as Alzheimer's and Prion diseases, e.g mad cow. In this context, it consists in the abnormal aggregation of proteins. Curves obtained by measuring the quantity of polymers formed in in vitro experiments are sigmoids: a long lag phase with almost no polymers followed by a fast consumption of all monomers. Furthermore, repeating the experiment under the same initial conditions leads to somewhat identical curves up to a translation.

The first study we did proposed a simplified stochastic model to analyze this phenomenon. For this model, when the volume gets large, the quantity of polymers has the typical sigmoidal shape. A second order result has also been obtained for this model. We were able to compute the asymptotic distribution of the lag time and express its variance. The parameters of the model have been obtained by using data given by Wei-Feng Xue, University of Kent.

The current project concerns a more sophisticated mathematical model. Indeed, we have added a conformation step: before polymerizing, proteins have to misfold. This step is very quick and remains at equilibrium during the whole process. Nevertheless, this equilibrium depends on the polymerization which follows the conformation step: this modelling leads to the study of averaging principles.

REGAL Project-Team

5. New Results

5.1. Highlights of the Year

- *Garbage collection for big data on large-memory NUMA machines*. We developed NumaGiC, a high-throughput garbage collector for big-data algorithms running on large-memory NUMA machines (see Section 4.1). This result, a collaboration with the Whisper team, will be presented at ASPLOS 2015 [29].
- *Explicit consistency*. We propose an alternative approach to the strong-vs.-weak consistency conundrum, *explicit consistency*. Static analysis identifies precisely what is the minimal amount of synchronisation that is necessary to maintain the invariants required by an application (see Section 5.3.11). This result will be presented at EuroSys 2015 [53].
- *Lower bounds and optimality for CRDTs*. This is the first paper to study the inherent lower bounds of replicated data types. The contribution includes derivation of lower bounds for several data types, improvement of some implementations, and proved optimality of others (see Section 5.3.10). This result was presented at POPL 2014 [25].

5.2. Distributed algorithms for dynamic networks

Participants: Luciana Bezerra Arantes [correspondent], Rudyar Cortes, Raluca Diaconu, Jonathan Lejeune, Olivier Marin, Sébastien Monnet, Franck Petit [correspondent], Karine Pires, Pierre Sens, Véronique Simon, Julien Sopena.

Nowadays, distributed systems are more and more heterogeneous and versatile. Computing units can join, leave or move inside a global infrastructure. These features require the implementation of dynamic systems, that is to say they can cope autonomously with changes in their structure in terms of physical facilities and software. It therefore becomes necessary to define, develop, and validate distributed algorithms able to managed such dynamic and large scale systems, for instance mobile *ad hoc* networks, (mobile) sensor networks, P2P systems, Cloud environments, robot networks, to quote only a few.

Efficiency in such environments requires specialised protocols, providing features such as fault or heterogeneity tolerance, scalability, quality of service, and self-*. 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.

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. Self-Stabilization.

We have also approached fault tolerance through self-stabilization. Self-stabilization is a versatile technique to design distributed algorithms that withstand transient faults.

In [43], we proposed a silent self-stabilizing leader election algorithm (SSLE, for short) for bidirectional connected identified networks of arbitrary topology. Starting from any arbitrary configuration, SSLE converges to a terminal configuration, where all processes know the ID of the leader, this latter being the process of minimum ID. Moreover, as in most of the solutions from the literature, a distributed spanning tree rooted at the leader is defined in the terminal configuration. This algorithm is written in the locally shared memory model. It assumes the distributed unfair daemon, the most general scheduling hypothesis of the model. Our algorithm requires no global knowledge on the network (such as an upper bound on the diameter or the number of processes, for example). We showed that its stabilization time is in $\Theta(n^3)$ steps in the worst case, where n is the number of processes. Its memory requirement is asymptotically optimal, *i.e.*, $\Theta(\log n)$ bits per processes. Its round complexity is of the same order of magnitude — *i.e.*, $\Theta(n)$ rounds — as the best existing algorithm designed with similar settings. To the best of our knowledge, this was the first self-stabilizing leader election algorithm for arbitrary identified networks that is proven to achieve a stabilization time polynomial in steps. By contrast, we show that the previous best existing algorithm designed with similar settings stabilizes in a non polynomial number of steps in the worst case.

We have also implemented SSLE in a high-level simulator to empirically evaluate its average performances. Experimental results tend to show that its worst case in terms of rounds ($\Theta(3n + D)$ rounds) is rare.

5.2.2. Dynamic Distributed Systems

The first key challenge in understanding highly dynamic networks consists in developing appropriate models that are as close as possible to the phenomena that one wishes to capture. This requires the use of a formalism sufficiently expressive to formulate complex temporal properties. Recently, a vast collection of concepts, formalisms, and models has been unified in a framework called Time-Varying Graphs (TVG) [0], which are represented as time-ordered sequences of graphs defined over a fixed set of nodes. A hierarchy of classes over TVG has been described, mainly depending on properties related to connectivity and recurrence of dynamic. Such an hierarchy is an interesting tool for study computability issues. As an example, if one is able to prove an impossibility result in a class of the hierarchy with strong properties, then this impossibility result also holds in any class of the hierarchy with (strictly) weaker properties. In this context, we provide a generic framework to prove impossibility results in this model [45]. This framework helps to formally prove classical arguments about convergence of sequence of time-varying graphs used to build counter-examples. We apply this generic framework to the study of covering problems (such as minimal dominating set and maximal matching) in the context of time-varying graphs. We obtain a characterization of the weakest topology assumption that makes these problems computable. We also propose a general time complexity measure since time-varying graph model lacks so far of such a definition.

5.2.3. Swarm of Mobile Robots

Swarm of autonomous mobile sensor devices (or, robots) recently emerged as an attractive issue in the study of dynamic distributed systems permits to assess the intrinsic difficulties of many fundamentals tasks, such as exploring or gathering in a discrete space. We consider autonomous robots that are endowed with visibility sensors (but that are otherwise unable to communicate) and motion actuators. Those robots must collaborate to solve a collective task, namely *exclusive perpetual exploration*, despite being limited with respect to input from the environment, asymmetry, memory, etc. The area to be explored is modeled as a graph and the exclusive perpetual exploration task requires every possible vertex to be visited infinitely often by every robot, with the additional constraint that no two robots may be present at the same node at the same time or may concurrently traverse the same edge of the graph.

In [28], we presented and implemented a generic method for obtaining all possible protocols for a swarm of mobile robots operating in a particular discrete space, namely an anonymous rings. Our method permits to discover new protocols that solve the problem, and to assess specific optimization criteria (such as individual coverage, visits frequency, etc.) that are met by those protocols. To our best knowledge, this was the first attempt to mechanize the discovery and fine-grained property testing of distributed mobile robot protocols.

[0] A. Casteigts, P. Flocchini, W. Quattrociocchi, and N. Santoro, Time-varying graphs and dynamic networks, International Journal of Parallel, Emergent and Distributed Systems 27(5):387-408, 2012

5.3. Management of distributed data

Participants: Pierpaolo Cincilla, Raluca Diaconu, Jonathan Lejeune, Mesaac Makpangou, Olivier Marin, Sébastien Monnet, Karine Pires, Dastagiri Reddy Malikireddy, Masoud Saeida Ardekani, Pierre Sens, Marc Shapiro, Véronique Simon, Julien Sopena, Vinh Tao Thanh, Serdar Tasiran, 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, located where, 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.
- How to apply our approaches towards addressing the above issues onto a challenging use case: achieving true scalability for online games.

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.

We have observed 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/scattered among the nodes, many source nodes may be available to heal the system, and thus, the system losses less pieces of data. In order to study data durability on a long term, we have designed a model, and implemented a discrete event based simulator that can simulate a 100 node system over years within several hours. Our model, SPLAD [49] (for scattering and placing data replicas to enhance long-term durability), allows us to vary the data scattering degree by tuning a selection range width. We are also studying the impact of the policy used while choosing a storing node within the selection range (e.g., randomly, the least loaded, or smarter policies like the power of two choices). This policy has an important impact on both the storage load distribution among nodes and the number of lost pieces of data.

5.3.2. Achieving scalability for online games

Massively Multiplayer Online Games (MMOGs) such as *World of Warcraft* constitute a great use case for the management of distributed data on a large scale. Commercial support systems for MMOGs rely almost exclusively on traditional client/server architectures that are centralized. These architectures do not scale properly, both in terms of the number of players and of the number objects used to model virtual universes that grow ever more complex. Most MMOGs avoid this problem by limiting the scale of the universe: the virtual environment is partitioned into several parallel and totally disconnected worlds, such as the *Realms* in *World of Warcraft*. Each partition, handled in a centralized way, limits the number of players it can host; avatars created on different partitions will never meet in the game.

From a systems point of view, achieving true scalability raises many challenging issues for MMOGs. For instance the system must be very reactive: if the update latency on a player node is too high, the game becomes unplayable. Since these games are meant to operate on a large scale, they induce a trade-off between availability and consistency of data. The consistency aspect is critical because MMOGs incur a high degree of cheating.

Designing and implementing a scalable service for Multiplayer Online Games requires an extensive knowledge of the habits, behaviors and expectations of the players. The first part of our work on MMOGs aimed at gathering and analyzing traces of real games offers to gain insight on these matters. We collected public data from a *League of Legends* server (information over more than 56 million game sessions): the resulting database is freely available online, and an ensuing publication [34] details the analysis and conclusions we draw from this data regarding the expected requirements for a scalable MMOG service.

We steered a second part of our work on MMOGs in 2014 towards designing a peer to peer refereeing system that remains highly efficient, even on a large scale, both in terms of performance and in terms of cheat prevention. Simulations show that such a system scales easily to more than 30,000 nodes while leaving less than 0.013% occurrences of cheating undetected on a mean total of 24,819,649 refereeing queries. This work got published in the Multimedia Systems Journal [21].

Finally, we also worked on the design of a scalable architecture for online games. The goal is to balance the load among nodes to allow the simulation of a whole, contiguous, virtual space.

5.3.3. Management of dynamic big data

Managing and processing Dynamic Big Data, where multiple sources produce new data continuously, is very complex. Static cluster- or grid-based solutions are prone to induce bottleneck problems, and are therefore ill-suited in this context. Our objective in this domain is to design and implement a Reliable Large Scale Distributed Framework for the Management and Processing of Dynamic Big Data. In 2014, we focused our research on data placement and on gathering traces from target applications in order to assess our future solutions.

With respect to placement, we worked on a scheme to store and access massive streams of data efficiently. We designed a solution that extends distributed prefix tree indexing structures for this purpose. Our new maintenance protocol anticipates every data insertion on provisional child nodes and thus significantly reduces overhead and improves query response time. This work has lead to the publication of an Inria research report (RR- 8637) [46].

With respect to application traces, we targeted sport tracker applications. Designing and implementing a big data service for sport tracker applications requires an extensive knowledge of both data distribution and input load. Gathering and analysing traces from a real world sports tracker service provides insight on these matters, but such services are very protective of their data due to competition as well as privacy issues. We avoided these issues by gathering public data from a popular sports tracker server called EndoMondo. The resulting database is freely available online, and allowed an in-depth analysis from a dynamic big data perspective. This study has lead to the publication of an Inria research report (RR- 8636) [47] .

5.3.4. 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 studied the popularity distribution and evolution of live video streams [31], [36].

5.3.5. Keyword-based Indexing and Search Substract for Structured P2P Information System

Number of large scale information systems rely on a DHT-based storage infrastructure. To help users to find suitable information, one attractive solution is to maintain an index that maps keywords to suitable data. Maintaining and exploiting an index distributed towards a DHT is confronted to the performance issue. Mainly, the computation of the intersection of postings related to provided keywords could generate too large traffic over the network; also one is confronted to some unbalanced on peers' load due to the fact that certain world are too popular!

In 2014, we propose *FreeCore*, a DHT-based distributed indexing substract that can be used to build efficient keyword-based search facilities for large scale information systems. A *FreeCore* index, considers keyword sets, then summarizes each set with a Bloom Filter. To limit the probability of false positive, we anticipate that one will use large size filters together enough hash functions. Thanks to this representation, we transform the searching problem, to the one of bitmaps matching as each query is also coded by a Bloom Filter. To distribute resulting summaries towards peers, *FreeCore* considers each summary as a sequence of binary keywords. Each binary keyword is assigned a peer and all summaries containing this binary keyword are stored at its assigned peer. Finally, to reduce the traffic overhead as well as the the size of local indices, *FreeCore* fragments each filter such as to factorize sequence of bits that occur more than once. In [40], we report the performances of the initial implementation of *FreeCore*. Thought a number of improvements were not included within this initial evaluation, *FreeCore* offers better performances than existing state of the art. Current work focusses on developping applications that exploit *FreeCore*.

5.3.6. Large-Scale File Systems

Storage architectures for large enterprises are evolving towards a hybrid cloud model, mixing private storage (pure SSD solutions, virtualization-on-premise) with cloud-based service provider infrastructures. Users will be able to both share data through the common cloud space, and to retain replicas in local storage. In this context we need to design data structures suitable for storage, access, update and consistency of massive amounts of data at the object, block or file system level.

Current designs consider only data structures (e.g., trees or B+-Trees) that are strongly consistent and partition-tolerant (CP). However, this means that they are not available when there is a network problem, and that replicating a CP index across sites is painful. The traditional approaches include locking, journaling and replaying of logs, snapshots and Merkle trees. All of these are difficult to scale using generic approaches, although it is possible to scale them in some specific instances. For instance, synchronization in a single direction (the Active/Passive model) is relatively simple but very limited. A multi-master (Active/Active) model, where updates are allowed at multiple replicas and synchronization occurs in both directions, is difficult to achieve with the above techniques.

Our previous work has shown that many storage indexing operations commute; this enables a the highly-scalable CRDT approach. For those that do not, the explicit consistency approach (Section 5.3.11) appears promising.

This work is part of a CIFRE agreeement with Scality (see Section 6.2.1).

5.3.7. 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.

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.

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.

An interesting side-effect of this research is an apples-to-apples comparison of many strong-consistency protocols. This work was published at LADIS 2014 [41] and at Middleware 2014 [33].

This research is supported in part by ConcoRDanT ANR project (Section 7.1.7) and by the FP7 grant SyncFree (Section 7.2.1.1).

5.3.8. 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. Both our simulations and the experimental results obtained with our prototype 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.

We have studied Gargamel's behavior while running over multiple geographically distant sites. One instance of Gargamel runs on each site, synchronizations among the different sites occur off the critical path [39]. Our experiments with the Amazon platform show that or solution can be used to support failures of whole sites.

5.3.9. 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).

CRDTs are the main topic of the ConcoRDanT ANR project (Section 7.1.7) and the FP7 grant SyncFree (Section 7.2.1.1). After developing the SwiftCloud extreme-scale CRDT platform (see Section 4.3), we are currently developing a flexible cloud database called Antidote (see Section 4.4).

5.3.10. Lower bounds and optimality of CRDTs

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, published at Principles of Programming Languages (POPL) 2014 [25]. Geographically distributed systems often rely on replicated eventually consistent data stores to achieve availability and performance. To resolve conflicting updates at different replicas, researchers and practitioners have proposed specialized consistency protocols, called replicated data types, that implement objects such as registers, counters, sets or lists. Reasoning about replicated data types has however not been on par with comparable work on abstract data types and concurrent data types, lacking specifications, correctness proofs, and optimality results. To fill in this gap, we propose a framework for specifying replicated data types using relations over events and verifying their implementations using replication-aware simulations. We apply it to seven existing implementations of 4 data types with nontrivial conflictresolution strategies and optimizations (last-writer-wins register, counter, multi-value register and observed-remove set). We also present a novel technique for obtaining lower bounds on the worst-case space overhead of data type implementations and use it to prove optimality of four implementations. Finally, we show how to specify consistency of replicated stores with multiple objects axiomatically, in analogy to prior work on weak memory models. Overall, our work provides foundational reasoning tools to support research on replicated eventually consistent stores.

5.3.11. Explicit Consistency: Strengthening Eventual Consistency to support application invariants

The designers of the replication protocols for geo-replicated storage systems have to choose between either supporting low latency, eventually consistent operations, or supporting strong consistency for ensuring application correctness. We propose an alternative consistency model, *explicit consistency*, that strengthens eventual consistency with a guarantee to preserve specific invariants defined by the applications. Given these application-specific invariants, a system that supports explicit consistency must identify which operations are unsafe under concurrent execution, and help programmers to select either violation-avoidance or invariant-repair techniques. We show how to achieve the former while allowing most of operations to complete locally, by relying on a reservation system that moves replica coordination off the critical path of operation execution. The latter, in turn, allow operations to execute without restriction, and restore invariants by applying a repair operation to the database state. We designed and evaluated Indigo, a middleware that provides Explicit

Consistency on top of a causally-consistent data store. Indigo guarantees strong application invariants while providing latency similar to an eventually consistent system.

This work was presented at W-PSDS 2014 [24] and LADIS 2014 [38]. It was selected for presentation at EuroSys 2015 [23]. This research is supported in part by the FP7 grant SyncFree (Section 7.2.1.1).

5.4. Memory management for big data

Participants: Antoine Blin, Lokesh Gidra, Sébastien Monnet, Marc Shapiro, Julien Sopena [correspondent], Gaël Thomas.

5.4.1. Garbage collection for big data on large-memory NUMA machines

On contemporary cache-coherent Non-Uniform Memory Access (ccNUMA) architectures, applications with a large memory footprint suffer from the cost of the garbage collector (GC), because, as the GC scans the reference graph, it makes many remote memory accesses, saturating the interconnect between memory nodes. We address this problem with NumaGiC, a GC with a mostly-distributed design. In order to maximise memory access locality during collection, a GC thread avoids accessing a different memory node, instead notifying a remote GC thread with a message; nonetheless, NumaGiC avoids the drawbacks of a pure distributed design, which tends to decrease parallelism. We compared NumaGiC with Parallel Scavenge and NAPS on two different ccNUMA architectures running on the Hotspot Java Virtual Machine of OpenJDK 7. On Spark and Neo4j, two industry-strength analytics applications, with heap sizes ranging from 160 GB to 350 GB, and on SPECjbb2013 and SPECjbb2005, NumaGiC improves overall performance by up to 45% over NAPS (up to 94% over Parallel Scavenge), and increases the performance of the collector itself by up to 3.6× over NAPS (up to 5.4× over Parallel Scavenge).

This research is accepted for presentation at the ASPLOS 2015 conference [29].

5.4.2. File cache pooling

Some applications, like online sales servers, intensively use disk I/Os. Their performance is tightly coupled with I/Os efficiency. To speed up I/Os, operating systems use free memory to offer caching mechanisms. Several I/O intensive applications may require a large cache to perform well. However, nowadays resources are virtualized. In clouds, for instance, virtual machines (VMs) offer both isolation and flexibility.This is the foundation of cloud elasticity, but it induces fragmentation of the physical resources, including memory. This fragmentation reduces the amount of available memory a VM can use for caching I/Os. We propose Puma [35] (for Pooling Unused Memory in Virtual Machines) which allows I/O intensive applications running on top of VMs to benefit of large caches.

This is realized by providing a remote caching mechanism that provides the ability for any VM to extend its cache using the memory of other VMs located either in the same or in a different host. Puma is a kernel level remote caching mechanism that is: (*i*) block device, file system and hypervisor agnostic; and (*ii*) efficient both locally and remotely. It can increase applications performance up to 3 times without impacting potential activity peaks.

WHISPER Team

6. New Results

6.1. Highlights of the Year

The paper "Faults in Linux 2.6" was published in the ACM journal Transactions on Computer Systems in June 2014 . It has been downloaded from the ACM digital library almost 300 times since then. The paper was reviewed in the Linux Weekly News, in the German professional IT website golem.de, and was the subject of an invited presentation at a joint session of the Linux Kernel Summit and LinuxCon North America.

Julia Lawall was invited to the 2014 Linux Kernel Summit, an invitation-only meeting of core Linux developers. She was subsequently invited to participate in the plenary Linux Kernel Developer Panel at LinuxCon Europe, with 2000 attendees.

Julia Lawall was invited to give a keynote at the conference Modularity (formerly AOSD) on her work on Coccinelle [16].

BEST PAPERS AWARDS :

[] **ACM Transactions on Computer Systems**. N. PALIX, G. THOMAS, S. SAHA, C. CALVÈS, G. MULLER, J. L. LAWALL.

6.2. Lock profiling in Java servers

Today, Java is regularly used to implement large multi-threaded server-class applications that use locks to protect access to shared data. However, understanding the impact of locks on the performance of a system is complex, and thus the use of locks can impede the progress of threads on configurations that were not anticipated by the developer, during specific phases of the execution. In our paper, "Continuously Measuring Critical Section Pressure with the Free-Lunch Profiler" [25], presented at OOPSLA 2014, we propose Free Lunch, a new lock profiler for Java application servers, specifically designed to identify, *in-vivo*, phases where the progress of the threads is impeded by a lock. Free Lunch is designed around a new metric, *critical section pressure* (CSP), which directly correlates the progress of the threads to each of the locks. Using Free Lunch, we have identified phases of high CSP, which were hidden with other lock profilers, in the distributed Cassandra NoSQL database and in several applications from the DaCapo 9.12, the SPECjvm2008 and the SPECjbb2005 benchmark suites. Our evaluation of Free Lunch shows that its overhead is never greater than 6%, making it suitable for *in-vivo* use.

6.3. Software engineering for infrastructure software

A kernel oops is an error report that logs the status of the Linux kernel at the time of a crash. Such a report can provide valuable first-hand information for a Linux kernel maintainer to conduct postmortem debugging. Recently, a repository has been created that systematically collects kernel oopses from Linux users. However, debugging based on only the information in a kernel oops is difficult. In a paper published at MSR [18], we consider the initial problem of finding the offending line, i.e., the line of source code that incurs the crash. For this, we propose a novel algorithm based on approximate sequence matching, as used in bioinformatics, to automatically pinpoint the offending line based on information about nearby machine-code instructions, as found in a kernel oops. Our algorithm achieves 92% accuracy compared to 26% for the traditional approach of using only the oops instruction pointer.

2014 was the second year of a two-year cooperation between Julia Lawall and David Lo of Singapore Management University, as part of the Merlion cooperation grant program of the Insitut Français. This cooperation resulted in four papers: two on word similarity [21], [26], one on bug localization [23], and one on an empirical study of testing practices in open source software [19]. As an offshoot of this work, Julia Lawall worked with the PhD student Ripon Saha of UT Austin and his advisors on the topic of assessing the effectiveness of a state-of-the-art bug localization technique on C programs as compared to Java programs [20]. This work built on the C parser developed for Coccinelle.

Finally, with colleagues from Aalborg University and with Nicolas Palix of Grenoble, Julia Lawall published an article in Science of Computer Programming assessing the applicability of Coccinelle to checking the coding style guidelines of the CERT C Secure Coding Standard [14].

6.4. Bugs in Linux 2.6

In August 2011, Linux entered its third decade. Ten years before, Chou et al. published a study of faults found by applying a static analyzer to Linux versions 1.0 through 2.4.1. A major result of their work was that the drivers directory contained up to 7 times more of certain kinds of faults than other directories. This result inspired numerous efforts on improving the reliability of driver code. Today, Linux is used in a wider range of environments, provides a wider range of services, and has adopted a new development and release model. What has been the impact of these changes on code quality? To answer this question, in an article published in ACM TOCS, we have transported Chou et al.'s experiments to all versions of Linux 2.6; released between 2003 and 2011. We find that Linux has more than doubled in size during this period, but the number of faults per line of code has been decreasing. Moreover, the fault rate of drivers is now below that of other directories, such as arch. These results can guide further development and research efforts for the decade to come. To allow updating these results as Linux evolves, we define our experimental protocol and make our checkers available.

6.5. Memory Monitoring in Smart Home gateways

Smart Home market players aim to deploy component-based and service-oriented applications from untrusted third party providers on a single OSGi execution environment. This creates the risk of resource abuse by buggy and malicious applications, which raises the need for resource monitoring mechanisms. Existing resource monitoring solutions either are too intrusive or fail to identify the relevant resource consumer in numerous multi-tenant situations. In our paper "Memory Monitoring in a Multi-tenant OSGi Execution Environment" [15], presented at CBSE 2014, we propose a system to monitor the memory consumed by each tenant, while allowing them to continue communicating directly to render services. We propose a solution based on a list of configurable resource accounting rules between tenants, which is far less intrusive than existing OSGi monitoring systems. We modified an experimental Java Virtual Machine in order to provide the memory monitoring features for the multi-tenant OSGi environment. Our evaluation of the memory monitoring mechanism on the DaCapo benchmarks shows an overhead below 46%. This work has been done as part of the PhD of Koutheir Attouchi [10] who was supported by a CIFRE grant with Orange Labs.

ALPAGE Project-Team

6. New Results

6.1. Highlights of the Year

Benoit Crabbé is a Junior Member of the Institut Universitaire de France (IUF) since October 2014. Two out of the five academic staff at Alpage are now member of the IUF, Laurence Danlos being a Senior Member since October 2013.

6.2. Automatic text normalisation

Participants: Benoît Sagot, Marion Baranes.

Since the emergence of the web, one of the goals of natural language processing (NLP) tools has been analysing raw noisy text documents such as blogs, review sites or social networks. These texts commonly contain misspellings, redundant punctuation, smileys, etc. Consequently they require specific preprocessing before being used in different NLP applications. That is why, we worked at Alpage on the development of a new corpora and the implementation of an automatic system for normalisation of such texts:

- **Corpus crap** In 2014, a large-scale extension of the number of normalisation rules used by the MElt part-of-speech tagger for processing noisy computer-generated content has been achieved. This work was carried out in the context of and based on corpora developed within the CoMeRe project, funded by the Institut de Linguistique Française and lead by Thierry Chanier [14].
- **Normalisation system** We have implemented a modular system which follows SxPipe [109]. This system detects if an unknown word to a reference lexicon corresponds to a non-word error (and is not a neologisme or a borrowing). Then, it attempts to normalize non-word errors and grammatical errors. In 2014, we focused on these two latter tasks. First, we have implemented a system which suggests one or several normalization candidates for these non-word errors. As described in [17], to do that, we use an analogy-based approach for acquiring normalisation rules and use them in the same way as lexical spelling correction rules. Secondly, we propose to normalize grammatical errors. To do that, we check for each word if it has common homophones. If this is the case, we consider these homophones as possible candidates for normalization. Finally, we filter all these candidates in order to keep only the one which is the most probable. This filtration is done using a probabilistic model based on a n-gram system. Moreover, the implementation of this system of normalisation motivated a side task. We developed an unsupervised method for acquiring pairs of lexical entries belonging to the same morphological family, i.e., derivationally related words, starting from a purely inflectional lexicon. This work, detailed in [16], allows us to create new linguistic resources for English, French, German and Spanish which contains derivational relations.

6.3. The impact of morphosyntactic processing on post-OCR error correction

Participants: Kata Gábor, Benoît Sagot, Pierre Magistry.

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. To our knowledge, very little work has been done to exploit linguistic analysis for post-OCR error correction. Within the PACTE project we are conducting research on reducing the OCR error rate by using contextual information and linguistic processing.

In 2014 we continued our investigations on how named entity recognition can benefit OCR error detection by applying context-aware error correction rules directly to the OCR output. Several grammars have been created or improved to adress OCR problems occurring within different types of named entities. As a result, the SxPipe-PACTE toolchain was created to correct named entities in a noisy input [45], [31].

While the symbolic error correction method works with a very high precision, its limitation lies in its relatively low coverage. In order to deal with the errors occurring outside the recognized entitites, we studied the possibility of using lattice-based part of speech tagging to select the best correction hypothesis in context. Different methods were investigated to generate correction hypotheses, using word alignment software or by observing frequently occurring error types. The initial results confirm that a significant number of the remaining OCR errors can be corrected via lattice-based tagging, as long as the noise introduced by correction hypotheses is controlled.

6.4. Linear-time discriminant syntactico-semantic parsing

Participants: Benoit Crabbé, Maximin Coavoux, Djamé Seddah.

In this module we study efficient and accurate models of statistical phrase structure parsing. We focus on linear time lexicalized parsing algorithms (shift reduce, left corner) with approximations entailing linear time processing. The existing prototype involves a global discriminant parsing model of the large margin family (Perceptron,Mira, SVM avatars) able to parse user defined structured input tokens [23]. Thus the model can take into account various sources of information for taking decisions such as word form, part of speech, morphology or semantic classes inter alia.

Our participation to the SPRML 2014 shared task on parsing morphologically rich languages has been a first step towards testing our model in a multilingual setting where we were among the state of the art systems and state of the art on some languages such as Polish. To our knowledge the parser is one of the fastest existing multilingual parser worldwide (4000 8000 tokens/sec.). In order to ease model design for multilingual settings, we currently study efficient feature selection procedures for automating model adaptation to new languages.

The ongoing investigation aims to integrate continuous semantic representations into the model such as word embeddings in order to leverage data sparsity and estimation issues recurrent in lexicalized parsing. To this end we study neural-network-based architectures for structured phrase structure parsing.

6.5. Playing with DyALog-based parsers

Participant: Éric Villemonte de La Clergerie.

Éric de la Clergerie has continued the development of DyALog-SR, a transition-based dependency parser running on top of DyALog and initiated in 2013 to participate to SPMRL'2013. Thanks to DyALog's tabulation functionalities, this parser implements a dynamic programming algorithm to explore larger search space through the use of beams.

In order to participate to SemEval'14 Task on "broad coverage semantic dependency parsing", DyALog-SR was extended to handle non-connected dependency graphs rather than standard dependency trees. This was achieved by considering a richer set of transitions, besides the usual Shift and Reduce transitions. However, while working, this extended set of transitions was not ensuring the expected gains when using beams. The issue was finally solved after long investigations, with the identification of multiple causes. One of them was related to the fact that transition paths of various lengths may lead to a final state. In consequence, a noop transition was added to compensate on shorter paths.

A second axe of work was a thorough use of DyALog-SR over the French TreeBank (FTB) to compare its performances to those published for other parsers. By enriching its set of features and improving the update strategy of the perceptron-based statistical model of DyALog-SR, we were able to reach state-of-the-art results.

However, the best results were obtained by coupling DyALog-SR with FRMG, our large-coverage French grammar (derived from a meta-grammar). The results from FRMG were used as features to guide the statistical DyALog-SR parser. This innovative step proved to provide us with the best results published so far for the FTB (over 90% of Labeled Attachement Score [LAS] over the test part of the FTB) [41].

The improvements of FRMG was pursued in 2014, at the level of the underlying meta-grammar (to extend its coverage over 96% on the FTB) but also by adapting the statistical models developed for DYALOG-SR (in replacement of older and slower SQLite-based models).

6.6. Multiword expressions and statistical parsing

Participants: Sarah Beniamine, Marie-Hélène Candito, Benoît Sagot, Djamé Seddah.

Multi-word expressions recognition (MWE recognition) and syntactic parsing are two tasks that have been extensively investigated. Yet, systems combining both tasks have been rather rare. In particulat, works on parsing have tended to use training and test data with gold MWEs (generally with each MWE) merged into one token. In 2013, Djamé Seddah led the organization of the first shared task on statistical parsing Morphologically Rich Languages (SPMRL) [127], hosted by the fourth SPMRL workshop. 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. The shared task proposed a data set for 9 languages. The French part of this data set is particular, in that it uses a representation combining MWEs and syntax, which allows to investigate techniques for performing parsing and MWE recognition. A first system was proposed for the dependency parsing track of the Shared Task, in collaboration with Matthieu Constant (LIGM, Université Marne-la-Vallée) [74]. This work investigates pipeline and joint architecture for both tasks. In 2014, Marie Candito and Matthieu Constant continued that line of work [2], focusing on using an alternative representation of syntactically regular MWEs, which captures their syntactic internal structure. The objective of such representation was two fold. First, it is well-known that the MWE status is not clear-cut, and that MWE status can hold due to syntactic and/or semantic criteria. In particular, syntactically regular MWEs exhibit various degrees of semantic non-compositionality. For such MWEs, an atomic representation fails to capture internal partial semantic composition, and also fails to take advantage of the internal syntactic regularity. Indeed, one hypothesis of this work was that augmenting the regularity of the syntactic representations could help parsing. The results of this work is that while this hypothesis could not be verified, the resulting system has comparable performance to that of previous works on this dataset, but it has the advantage of predicting both syntactic dependencies and the internal structure of MWEs, a crucial feature to capture the various degrees of semantic compositionality of MWEs.

In the same time, Sarah Beniamine and Benoît Sagot also investigated the use of internal regular structures for MWEs, yet for *syntagmatic* syntactic parsing. The objective is to guide a parser with predicted MWEs, while keeping a regular syntactic representation.

6.7. Graph-based approaches for deep-syntatic and semantic parsing

Participants: Corentin Ribeyre, Djamé Seddah, Éric Villemonte de La Clergerie.

With most state-of-the-art statistical parsers routinely crossing a ninety percent performance plateau in capturing tree structures, the question of *what next* crucially arises. Most of the structures used to train current parsing models are degraded versions of a more informative data set: the Wall Street journal section of the Penn treebank ([91]) which is often stripped from its richer set of annotations (i.e. traces and functional labels are removed), while, for reasons of efficiency and availability, projective dependency trees are often given preference over richer graph structures [96], [107]. This led to the emergence of *surface* syntax-based parsers [70], [97], [100] whose output cannot by itself be used to extract full-fledged predicate argument-structures. For example, control verb constructions, it-cleft structures, argument sharing in ellipsis coordination, etc. are among the phenomena requiring a graph to be properly accounted for. The dichotomy between what can usually be parsed with high accuracy and what lies in the deeper syntactic description has initiated a line of research devoted to closing the gap between surface syntax and richer structures.

At Alpage, we built our work on the widely known transition-based parsing approach [95], which is state-of-the-art to parse surfacic syntatic trees [141]. Shift-reduce transition-based parsers essentially rely on *configurations* formed of a stack and a buffer, with stack transitions used to move from a configuration to the next one, until reaching a final configuration.

6.8. English Broad-coverage Semantic Dependency Parsing

Participants: Corentin Ribeyre, Djamé Seddah, Éric Villemonte de La Clergerie.

We successfully tested our graph-based approach described in Section 6.7 on a shared task on broad-coverage semantic dependency parsing part of the International Workshop on Semantic Evaluation (SemEval 2014, [99]). We were given three resources, which constitute parallel semantic annotations over the same common text (the Penn Treebank (PTB), [91]). The first one is part of the tectogrammatical layer of the Prague Czech-English Dependency Treebank, the second one is the reduction of the Minimal Recursion Semantics, available through the HPSG annotation of the PTB, into bi-lexical dependencies [82]. Finally, the third one is the predicate-argument structures extracted from the Enju Parser [131]. The shared task consisted of two tracks: a closed one where we needed to use these three resources only and an open one, where we could use whatever we needed to produce the best semantic representations.

At Alpage, we developed two semantic parsers: The first one is based on a previous work on DAG parsing [107] and the second one on the FRMG surfacic syntactic parser [133]. We use two parsers to assess the validity of our approach. The top performing models we submitted used a mix of syntactic features (tree fragments from a constituent syntactic parser [100], dependencies from a syntactic parser [58], elementary spinal trees using a spine grammar [126], etc.) to improve our results. Our intuition is that syntax and semantic are not independent of each other and using syntax could improve semantic parsing. Our systems performs well and were able to compete with the top performers. Those systems, as well as the software needed to parse these new data sets, are already available.

6.9. Development of syntactic and deep-syntactic treebanks: Extending our Coverage

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

Taking its roots in the teams that initiated the first syntactically annotated the French Treebank, the first metagrammar compiler and one of the best wide coverage grammars, Alpage has a strong tendency to focus on creating pioneer resources that serve both to extend our linguistics knowledge and to nurture accurate parsing models. Recently, we focused on extending the lexical coverage of our parsers using semi-supervized techniques (see above) built on edited texts. In order to evaluate these models, we built the first free out-domain treebank for French (the Sequoia treebank, [69]) covering various domains such as Wikipedia, Europarl and bio medical texts on which we established the state-of-the-art. Exploring other kind of texts (speech, user generated content), we faced however various issues inherently tied to the nature of these productions. Syntactic divergences from the norm are actually prominent and are a severe bottleneck for any data driven parsing model. Simply because a structure not present in a training set cannot be reproduced. This analysis naturally occurred as a side effect of our experiments in parsing social media texts. Actually, the first version of the French Social Media Bank (FSMB) was conceived as a stress test for our tool chains (tokenization, tagging, parsing). Our recent experiments showed that to reach a decent performance plateau, we need to include some of the target data into our training set. Focusing on processing direct questions and social media texts, we built two treebanks of about 2,500 sentences each: one devoted to questions and one built to extend the FSMB [0]. These initatives are funded by the Labex EFL.

- The French Social Media Bank 2.0: We are about to release the second part of the FSMB, 2600 sentences from Twitter, Facebook and other sources, with an extended annotation scheme able to describe more precisely the various phenomena at stakes in the social media text streams. To do so we extended our pre-processing chain (included and available in the MeLT tagger) to include a much more robust normalizer and tokenizer than the one we used to build the first version of the FSMB. The building phase being over, publications on this topics are on preparation.

[0] Let us note that the ever evolving nature of user generated content makes this a necessity.

- The French Question Bank: The building of a treebank made solely of questions comes from the simple fact that in both the FTB and the Sequoia treebank, there's only 150 direct questions. Making the parsing of such constructions extremely difficult for our data driven parsers. Following our now classical methodology, we selected more than 3200 sentences coming from governmental sources, from the TREC ressources – allowing to have a strong set of aligned sentences with the English ressources – and from social media sources as well. In the case of the TREC part, those are the questions used by [85], which allows some potentially interesting cross-language experiments. Unlike in the English Question Bank, phrasal-movement are annotated with functional paths and not traces. This allows to maintain a strong compatibility with the FTB annotation scheme. Our Question bank is the only resources of its kind for any other languages than English.

Both ressources are available in constituency and dependency. The later being still verified for the FSMB 2.0.

Note that we just started another annotation campaign aiming at adding a deep syntax layer to these two data sets, following the Deep Sequoia as presented above. These resources will prove invaluable to building a robust data driven syntax to semantic interface.

In the same time, Alpage collaborated with the Nancy-based Inria team Sémagramme in the domain of deep syntax analysis. Deep Syntax is intended as an intermediary level of representation, at the interface between syntax and semantics, which partly abtracts away from syntactic variation, and aims at providing the canonical grammatical functions of predicates. This means for instance neutralizing diathesis alternation and making explicit argument sharing, such as occurring for infinitival verbs. The advantage of a deep syntactic representation is to provide a more regular representation to serve as basis for semantic analysis. Note though it is computationally more complex, as we switch from surface syntactic trees to deep syntactic graphs, since shared arguments are made explicit.

We collaboratively defined a deep syntactic representation scheme for French and built a gold deep syntactic treebank [21], [43]. More precisely, each team used an automatic surface-to-deep syntax converter module, applied it on the Sequoia corpus (already annotated for surface syntax), and manually corrected it. Remaining differences were collaboratively adjudicated. The surface-to-deep syntax converter tool used by Alpage is built around the OGRE Graph Rewriting Engine built by Corentin Ribeyre [105].

The Deep Sequoia Treebank is too small to train a deep syntactic analyzer directly. In order to obtain more annotated data, we further used the surface-to-deep syntax converter to obtain predicted (non validated) deep syntactic representations for the French Treebank [36], which is much bigger than the Sequoia treebank (more than 18.000 sentences compared to $3,000$ sentences). We performed an evaluation of a small subset of the resulting deep syntactic graphs. The high level of performance we obtained (more than 98% of F-score in labeled dependencies recovery task) which suggests that the deep syntax version of the French Treebank can be used as pseudo-gold data to train deep syntactic parsers, or to extract syntactic lexicons augmented with quantitative information.

6.10. Towards a French FrameNet

Participants: Marie-Hélène Candito, Marianne Djemaa, Benoît Sagot.

The ASFALDA project [0] is an ANR project coordinated by Marie Candito. 5 partners collaborate on the project, on top of Alpage : the Laboratoire d'Informatique Fondamentale de Marseille(LIF), the Laboratoire de Linguistique Formelle (LLF), the MELODI team (IRIT - Toulouse) and the CEA-List. It 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 [57], [0] which provides a structured set of prototypical situations, called *frames*, along with a semantic characterization of the participants of these situations (called *frame elements*). The resulting resources will consist of :

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

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 2014, we first finished the work on the lexicon, which was started in 2013 [19]. The step 2 (semantic annotations on top of syntactic representations) is ongoing :

- We wrote the annotation guide. In particular Marianne Djemaa focused on how to annotate phenomenon known to exhibit syntax/semantic divergences [42].
- We designed the annotation workflow and built an automatic pre-annotator, which proposes candidate semantic annotations that must be disambiguated manually.
- We started in july 2014 to manage six annotators, who were hired to perform the manual annotation phase.

6.11. Towards a morpho-semantic resource for French designed for Word Sense Disambiguation

Participant: Lucie Barque.

The most promising WSD methods are those relying on external knowledge resources [93] but semantic resources for French are scarse. Moreover, existing resources offer fine grained sense dictinctions that do not fit to WSD. Our aim is to provide the NLP community with a broad-coverage morpho-semantic lexicon for French that relies on coarse-grained sense distinctions for polysemic units. Preliminary results concern nouns, on which we have first focused because their semantic description, compared to verbs, crucially lacks (for information retrieval, for instance) and because the regular polysemy phenomenon (recurring cases of polysemy within semantic classes) mainly occurs in nominal semantic classes:

- We proposed a linguistically motivated description of general semantic labels for nouns, that will allow for coarse-grained sense distinctions [40]
- Regular polysemy of nouns that can denote an event or a participant of this event has also been described for a large number of French nouns in [12]

6.12. Development of Verb∋net

Participants: Laurence Danlos, Quentin Pradet.

VerbNet is an English lexical resources for verbs, which is internationally known and widely used in numerous NLP applications [89]. Verb∋net is a French adaptation of this resource. It is semi-automatically developed thanks to the use of two French existing resources created in the 70's: LG, Lexique-Grammaire developed at LADL under the supervision of Maurice Gross, and LVF, Lexique des verbs du français by Dubois and Dubois-Charlier. The idea is to map English classes, which gather verbs with a common syntactic and semantic behavior, into classes of LG and LVF, then to manually adapt the syntactic frames according to French grammar while keeping the thematic roles and the semantic information, [35], [28]. This work is currently under progress in collaboration with Takuya Nakamura (Institut Gaspard Monge) and the resource should be freely available in 2015.

6.13. Development of FDTB1

Participants: Laurence Danlos, Margot Colinet, Jacques Steinlin.

FDTB1 is the first step towards the creation of the French Discourse Tree Bank (FDTB) with a discourse layer on top of the syntactic one which is available in the French Tree Bank (FTB). In this first step, we have identified all the words or phrases in the corpus that are used as "discourse connectives". The methodology was the following: first, we highlighted all the items in the corpus that are recorded in LexConn [106], a lexicon of French connectives with 350 items, next we eliminated some of these items with the following criteria:

1. first, we filtered out the LexConn items that are annotated in FTB with parts of speech incompatible with a connective use, e.g. *bref* annotated as *Adj* instead of *Adv*, *en fait* annotated as *Pro V* instead of (compound) *Adv*;
2. second, as we lay down for theoretical and pratical reasons that elementary arguments of connectives must be clauses or VPs, we filtered out e.g. LexConn prepositions that introduce NPs;
3. last, we filtered out LexConn prepositions and adverbials with a non-discursive function.

The last criterion requires a manual work contrarily to the two others. For example the preposition *pour (to)*, is ambiguous between a connective use *(Fred s'est dépeché pour être à la gare à 17h (Fred hurried to be at the station at 17h))* and a preposition introducing a complement *(Fred s'est dépeché pour aller à la gare (Fred hurried to go to the station))*, and the disambiguation between the two uses is subtle and so the topic of a long paper [22], whose results have been used to enhance Le*fff*, [44].

The FDTB corpus contains 18 535 sentences and FDTB1 identifies 9 833 discourse connectives. This ressource is freely available.

6.14. Discourse Parsing

Participants: Laurence Danlos, Chloé Braud.

Discourse parsing goal is to reflect the rhetorical structure of a document, how pieces of text are linked in order to form a coherent document. Understanding such links could benefits to several other natural language applications (summarization, language generation, information extraction...). A discourse parser corresponds to two major subtasks: a segmentation step wherein discourse units (DUs) are extracted, and a parsing step wherein these DUs are (recursively) related through "discourse (rhetorical) relations". The more difficult task in discourse parsing is the labeling of the relations between DUs, especially when no so-called connective overtly marks the relation (we then talk about implicit relations as opposed to explicit ones). In her PhD work, Chloé Braud develops a discourse relation classifier, carrying experiments on French and English. Focusing on the problem on implicit relation identification, this work tries to tackle the lack of manually annotated data, a discourse specific difficulty, by exploiting the similarities between explicit and implicit relations. In 2014, this work lead to systems based on domain adaptation methods [18], [13], demonstrating improvements on the French corpus Annodis [56].

6.15. Multilingual and cross-lingual terminology extraction

Participants: Valérie Hanoka, Benoît Sagot.

Language diversity spans more than 7000 languages. Among them, 24 macrolanguages [0] have at least 50 million first-language speakers. Traditional terminology techniques, which are mostly based on language-dependent linguistic tools (part of speech tagging, phrase chunking) requires a considerable effort to be developed for a new language. This effort is likely to be even more critical if the term extraction is to be based on noisy text (i.e. displaying linguistic creativity, spelling errors and ungrammatical sentences). In this context, the need has arisen to examine the issue of a less language-specific method for term extraction.

To that end, our approach take advantage of existing language typologies in order to alleviate for the lack of language-dependent linguistic processing. We based our reflexions and experiments on a sample of 7 typologically different language: Arabic, Chinese, English, French, German, Polish and Turkish.

[0] A macrolanguage is defined as "multiple, closely related individual languages that are deemed in some usage contexts to be a single language" in the ISO 639-3 standard.

As a starting point, we considered the minimal textual preprocessing (character normalization, segmentation) needed to allow for a comprehensive multilingual approach to automatic term extraction. In order to gain further insight on the influence of the morphology for term extraction, we examined the impact of the deletion of selected morphological information on words of morphologically rich languages.
For the different settings, models based on Conditional Random Fields (CRF) have been trained on existing gold data. We proposed an adapted version of the evaluation algorithm of [94] able to issue terminological scores for all the language of our sample. The scores thus obtained allowed to identify the best experimental setting for each language tested.

The results were surprising in two ways: First, the cross-lingual [0] application of models works well (the best cross-lingual models' accuracies range from 0.8% to 0.97%). Secondly, the languages which makes the overall best cross-lingual models are those who have the richest morphology (i.e: Turkish).

Finally, we developped and used a multilingual translation graph [32] to extend the multilingual terminology obtained using two methods: those presented in [83] and a more formal one, based on a simulated annealing clustering algorithm.

6.16. Word order variation in Old French

Participants: Benoit Crabbé, Alexandra Simonenko, Benoît Sagot.

As participant of the strand *Experimental Grammar* of the Labex EFL project *Empirical Foundations of Linguistics*[0] we study word order issues on Old French and more specifically the relative ordering of complements of ditransitive verbs. The inquiry seeks to identify several factors influencing the ordering of Old French complementation in different texts (varying in dates and genres) by carrying quantitative and statistical work from annotated Old French data. [0]

The quantitative results will be compared with what is known from corpus studies on the relative ordering of subject and complement in Old French [90]. It will also be compared to the quantitative results obtained on the relative ordering of complements of ditransitive verbs in Modern French [8] and modern English [64]. This comparative perspective is expected to provide new insights on French language evolution.

6.17. Cross linguistic factors governing word order

Participants: Benoit Crabbé, Kristina Gulordava.

In many languages, flexible word order often has a pragmatic role and marks the introduction of new information, a focus or a topic shift. Other cases of language- internal word order variation are alternations between two options such as *Mary gave John a book* and *Mary gave a book to John*, which are conditioned on syntactic and semantic factors such as the complexity of the constituents (as in *Mary gave John a book she had read ten times*), their animacy or the meaning of the verb [63].

One of the goals of this module is to investigate the connection between the quantitative aspects of word order variation across languages and the quantitative aspects of word order variation within a language. We study the corresponding patterns in language-internal variation by looking at the syntactically annotated corpora of various languages. Focusing on the variation of the internal word order of the noun-phrase as a case study, we explore to which extent a computational corpus-based analysis can provide new evidence not only for empirical, but also for theoretical linguistic research.

6.18. Anaphoricity detection and coreference resolution

Participant: Emmanuel Lassalle.

[0] A model trained on data of one language and applied to data of another language
[0] www.labex-efl.org
[0] SRCMF corpus: http://srcmf.org/; MCVF: http://www.voies.uottawa.ca

Resolving coreference in a text, that is, partitioning mentions (noun phrases, verbs, etc) into referential entities, is a challenging task in NLP leading to many different approaches. Anaphoricity detection, on the other hand, consists in deciding whether a mention is anaphoric (aka discourse-old) or non-anaphoric (discourse-new). This task is strongly related to coreference resolution and has been mainly addressed as a preliminary task to solve, leading to pipeline architectures.

A first line of work compares several methods for learning latent structures encoding coreference clusters that optionally take into account very accurate constraints on mention pairs. We study the relationship between standard decoding strategies used with pairwise models and those used with structured learning of latent structures, providing both topological and empirical comparisons. We also show that further gains can be obtained by the addition of pairwise constraints. Our experiments on the CoNLL-2012 dataset show that our best system obtains state-of-the-art results, and significant gains compared to standard locally-trained models.

Our second line of work introduces a new structured model for learning anaphoricity detection and coreference resolution in a jointly. Specifically, we use a latent tree to represent the full coreference and anaphoric structure of a document at a global level, and we jointly learn the parameters of the two models using a version of the structured perceptron algorithm. This model is refined by the use of pairwise constraints, and our experiments on the CoNLL-2012 English datasets show large improvements in both coreference resolution and anaphoricity detection, compared to various competing architectures. Our best coreference system obtains a CoNLL score of 81.97 on gold mentions, which is to date the best score reported on this setting.

This work has been achieved in collaboration with Pascal Denis, a former Alpage member, now at Inria Lille-Nord-Europe (EPI Magnet).

RITS Team

6. New Results

6.1. Highlights of the Year

YoGoKo [0], a startup company of RITS, was founded in 2014 by employees from three research institutes: Mines ParisTech, Telecom Bretagne and Inria. YoGoKo makes use of softwares developed in teams specialized in Internet technologies. RSM (Telecom Bretagne), CAOR (Mines ParisTech) and RITS (Inria) are research teams have been working together since 2006 on innovative communication solutions applied to Intelligent Transportation Systems. They contributed to several collaborative R& D projects related to ITS (CVIS, ITSSv6, GeoNet, DriveC2X, SCORE@F, ...). In 2012, these laboratories engaged together into the development of a common demonstration platform which comprises connected vehicles (fleet of conventional vehicles from Mines ParisTech and fleet of autonomous vehicles from Inria), roadside equipments and cloud-based services. YoGoKo demonstration platform was finally revealed on Feb. 11 th 2014 during the Mobility2.0 event organized by the French Ministry of Transport. This successful demonstation and the extremely warmfull feedack gained at this occasion triggered the launch of YoGoKo as a company. YoGoKo develops innovative communication solutions for fixed and mobile multi-connected devices. The objective is to maintain secure and continuous connectivity with their communication peers, either in their immediate environment or a remote location (control centers or Internet hosts).

6.2. Development of a Platform for Arbitration and Sharing Control Applications

Participants: David Gonzalez Bautista, Vicente Milanes Montero, Fawzi Nashashibi, Joshué Pérez Rastelli.

RITS have been leading the activities in the framework DESERVE project, related to arbitration and control sharing in automated vehicle. The analysis of existing vehicle control (and arbitration) solutions, considering the driver in the control loop is the main challenge of this work. We consider sharing control techniques and different solutions in the task management. New standard in the taxonomy of autonomous driving, as the SAE J3016, are considered in the arbitration and sharing control design. The aim is to allow the applications to make effective use of the driver model to improve the acceptability of the functions developed, as: Driver Drowsiness and Driver intention.

The arbitration module is defined into the IWI manager of the DESERVE abstraction. This component determines the action to be taken by the driver. The Driver Assistance Systems involve two main decision makers: the driver and the automated systems. This module considers different inputs, as follow: the Trajectory planning, Driver stage, Risk Management. The output determines who should take the control of the vehicle and the level of arbitration (or disposal) of the driver in different situations. This work uses the software tool FEMOT (Fuzzy Embedded MOTor). More detail can be found in [46].

6.3. Optimal Energy Consumption for Urban Electric Vehicles

Participants: David Gonzalez Bautista, Vicente Milanes Montero, Joshué Pérez Rastelli.

[0] http://www.yogoko.fr/

RITS team is specially supporting two kinds of transport systems: electric mass-produced vehicles and Cybernetic Transport Systems (also electrically propelled) for urban environments. One of the key factors for getting a higher market penetration of such vehicles is their autonomy. Having this in mind, the goal of this research line is to create optimal algorithms for improving electric vehicles' battery life. It covers two specific arenas: 1) determining optimal path planning in terms of energy saving (proposed for 2015); and 2) once the route is determined, generating an adequate speed profile for covering that path. The latter objective has been investigated during 2014. Energetic model of vehicle dynamics have been developed in order to determine the lowest consumption for each of the route segments. It has permitted to develop speed references between segments combined with polynomial transition functions for the whole route to be covered. Additionally, a high-level fuzzy controller has been also designed to make the system robust to low-level failures on reference tracking. Up to 20% of battery savings have been obtained in the first tests with the proposed algorithm, showing the proper performance of the system. Additional work for adding more information from the environment as other road agents or potential unexpected diversion on the road will be also investigated during 2015 for adapting the algorithm to more realistic environments. This work has been also developed in cooperation with MSc students from Simon Bolivar University (Venezuela) and AGMUS University System (Puerto Rico, US).

6.4. Perception and control strategies for autonomous docking for electric freight vehicles

Participants: Joshué Pérez Rastelli, Evangeline Pollard, Vicente Milanes Montero, Fawzi Nashashibi.

The freight transportation is defined as the process of carrying goods and persons from one given point to another. Recently, urban freight transportations have been used as an alternative for the delivery problems of goods in urban environments. The present work is developed in the framework of the Furbot project (FP7), which presents a solution for future urban freight transport with new light-duty architecture with full-electrical vehicles. We focused on the onboard intelligent units, dedicated to improve the perception and control systems onboard the vehicle for the parking/docking process, considering loading and unloading phases of the freight transport procedure. Two lasers were placed on the vehicle in order to localize it with respect to the freight box. A polynomial approach is used for the trajectory planning for a smooth docking maneuver. This proposal was first tested in a 3D simulator, and then validated in a real platform. The results presented in [45] shows the good behavior of our approach, which will be implemented in the FURBOT vehicle at the end of the project.

6.5. Description and technical specification of Cybernetic Transportation Systems: an urban transportation concept

Participants: Joshué Pérez Rastelli, Vicente Milanes Montero, David Gonzalez Bautista, Armand Yvet, Fawzi Nashashibi.

The Cybernetic Transportation Systems (CTS) is an urban transportation concept based on two ideas: the car sharing and the automation of dedicated systems with door-to-door capabilities. In the last decade, many European projects have been developed in this context, where some of the most important are: Cybercars, Cybercars II, CyberMove, CyberC3 and CityMobil, where a first fleet of vehicles were developed by different companies and research centers around Europe, Asia and America. Considering these previous works, the FP7 project Citymobil II is in progress since 2012. Its goal is to solve some of the limitations found so far, including the definition of the legal framework for autonomous vehicles on urban environment. Much of the perception and control software has been improved in the Inria's Cybus. New guidance functionalities were developed, mainly with the introduction of stereovision-based SLAM, and Bezier curve in path planning generation. In this work, automated CTSs involved are used in the different showcases in European cities. This work presents the different improvements, adaptation and instrumentation of the vehicle used. Results show tests in our facilities at Inria-Rocquencourt (France) and the first showcase at León (Spain).

6.6. Evidential Simultaneous Localization And Mapping to describe intersection

Participants: Guillaume Trehard, Evangeline Pollard, Fawzi Nashashibi.

Intersections management remains a tough challenge to tackle before reaching autonomous driving in urban environment. The field of view of the vehicle is often limited by several sensors occlusions, the shapes and priority rules can significantly differ from an intersection to another and road users from pedestrian to public transports have to cross each other in sometimes complex manners.
In this context, mapping the surrounding of the vehicle and being able to estimate its position regarding a global database is crucial.
A solution of *Simultaneous Localization And Mapping* (SLAM) have then been proposed based on a 2D LIDAR sensor [49]. In the rich SLAM literature, the originality of this method lays in the use of Transferable Belief Model (TBM) framework instead of a classic probabilistic one. If this proposition was just a change of mathematical context, TBM led to an explicit management of not-known and conflict information so that its application to SLAM algorithm appeared to be really effective and robust in crowded situations. The proposed solution indeed enables to provide a map of the *static* environment crossed by the vehicle and to detect mobile obstacles in the same process and without additional tracking system.
This *Evidential SLAM* have then been tested with success over different sequences and laser set-ups extracted from the KITTI database [50].
Researches are now focused on the fusion between this SLAM solution and a Global Navigation Satellite System (GNSS) receiver to enable a map-matching on a database such as Open Street Map.

6.7. Laser based road obstacle classification

Participants: Pierre Merdrignac, Evangeline Pollard, Oyunchimeg Shagdar, Fawzi Nashashibi.

Vehicle and pedestrian collisions often result in fatality to the vulnerable road users, indicating a strong need of technologies to protect such vulnerable road users. Laser sensors have been extensively used for moving obstacles detection and tracking. Laser impacts are produced by reflection on these obstacles which suggest that more information is available for their classification. This year, we introduced the design of a new system for road obstacles classification that is divided in four parts: definition of geometric features, selection of the best features, multi-class *segment* classification based on Support Vector Machines (SVM) and *track* classification from SVM decision values integration. Our study discloses a sorted list of useful features for road obstacle recognition that were used to construct a multi-class SVM. Finally, we tested our system with 2D and 3D laser sequences and shown that it can successfully estimate the class of some road obstacles around the vehicle.

6.8. Deformable Parts Model based approach for on-road object detection and classification

Participants: Wei-Lin Ku, Evangeline Pollard, Anne Verroust-Blondet.

An important perception problem for driver assistance is the detection of the road obstacles and the recognition of their type (cars, cycles, pedestrians). This year, we tackled the on-road objects detection problem by testing and improving vision-based methods. We proposed and compared several DPM based strategies for on-road object detection and classification, laying emphasis on the problem of detecting smaller/occluded cars and pedestrians. A hybrid approach combining detection from small/large models trained with different clustering method has been introduced to boost the detection performance in both Average Precision and Maximum Recall in every difficulty level. Finally, a geometry reasoning based filtering has been employed to eliminate false alarms while preserving a great deal amount of true positives. Experimental results showed the improvement both in hybrid and geometry reasoning approaches. Most of this work has been done during the internship of Wei-Lin Ku.

6.9. Saturated Feedback Control for an Automated Parallel Parking Assist System

Participants: Mohamed Marouf, Fawzi Nashashibi, Plamen Petrov.

In 2014, RITS extended its activities in the design and development of specific automated maneuvres. One particular interesting topic is the parallel parking problem of automatic front-wheel steering vehicles. The problem of stabilizing the vehicle at desired position and orientation is seen as an extension of the tracking problem. A saturated control has been proposed which achieves quick steering of the system near the desired position of the parking spot with desired orientation and can be successfully used in solving parking problems. In addition, in order to obtain larger area of the starting positions of the vehicle with respect to the parking spot for the first reverse maneuver of the parallel parking, an approach using saturated control with two different levels of saturation is proposed. The vehicle can be automatically parked by using one or multiple maneuvers, depending on the size of the parking spot. Simulation results were presented first in [44] to confirm the effectiveness of the proposed control schemes. New results extended to all types of parking lots shapes were recently obtained using this approach. The validation has been performed with real vehicles in the Inria test site.

6.10. 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 question whether the Wi-Fi-based V2P communications meet the requirements of the pedestrian safety. This year, we studied the performances of the V2P communications especially for the 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. This work has been published in [34].

6.11. Multicast Communications for Cooperative Vehicular Systems

Participants: Ines Ben Jemaa, Oyunchimeg Shagdar, Arnaud de La Fortelle.

Vehicular communications allow emerging new multicast applications such as fleet management and point of interest (POI). Both applications require Internet-to-vehicle multicasting. These approaches could not be applied to vehicular networks (VANET) due to their dynamic and distributed nature. In order to enable such multicasting, our work deals with two aspects. First, reachability of the moving vehicles to the multicast service and second, multicast message dissemination in VANET. We propose a simplified approach that extends Mobile IP and Proxy Mobile IP. This approach aims at optimizing message exchange between vehicles and entities responsible for managing their mobility in Internet. To study the dissemination mechanisms that are suitable for fleet management applications, we propose to revisit traditional multicast routing techniques that rely on a tree structure. For this purpose, we study their application to vehicular networks. In particular, as vehicular networks are known to have changing topology, we study the application of Multicast Adhoc On Demand Vector, MAODV. We propose then Motion-MAODV [35] [16] an improved version of MAODV that aims at enhancing routes built by MAODV in vehicular networks and guarantee longer route lifetime. Finally, to enable geographic dissemination as required by POI applications, we propose the routing protocol Melody that provides a geocast dissemination in urban environments. Through simulations, Melody ensures more reliable and efficient packet delivery to a given geographic area compared to traditional geo-broadcasting schemes in highly dense scenarios.

6.12. Visible Light Communication for ITS applications

Participants: Mohammad Abu Alhoul, Oyunchimeg Shagdar, Fawzi Nashashibi.

Visible Light Communication (VLC) technology is an efficient supportive communication technology for platooning applications over short inter-vehicle distances. After implementing a complete VLC channel model, which enabling precise calculations of the optical link performance for different intervehicle distances presented in our previous work [1], this year we have studied and proposed tracking-alike method aiming at ensuring the continuity of the Line-of-Sight (LOS) and extending the Field of view (FOV) limitations. This method benefits from the exchanged information about the relative directional position of each member of the platoon, together with front and rear facing directions of each vehicle, which can be very useful data for building a reliable smooth geometrical-based compensation method. The simulation results showed that trajectory influences on the optical incidence and irradiance angles can be compensated efficiently and without deploying any tracking method.

6.13. Study on the IEEE 802.11p Channel Congestion Problem

Participant: Oyunchimeg Shagdar.

The IEEE 802.11p is a standardized WiFi technology dedicated to V2X communications for especially road safety and efficiency applications. It is expected that vehicles periodically broadcast messages to announce their existences using the IEEE 802.11p frequency channel. However, because the IEEE 802.11p has a limited wireless bandwidth, in dense traffic conditions the V2V communications performances are poor, failing to satisfy the application requirements. In RITS, we study the issue and develop congestion control algorithms. In 2014, we studied the reactive distributed congestion control algorithm, proposed by the European Telecommunications Standards Institute (ETSI), and showed that the algorithm creates unstable resource utilization, which can cause the reactive Distributed Congestion Control (DCC) to perform worse than non-DCC systems. We proposed an asynchronous algorithm, where DCC control is made in such a way that channel resource is used in an asynchronous manner by the different stations. Our results show that the asynchronous DCC approach outperforms both the non-DCC and reactive DCC mechanisms. The work has been reported at a ETSI meeting in December 18, 2014 [54].

6.14. Study on V2V Communications and Emergent Behavior of Heuristically-Driven Intelligent Vehicles

Participant: Oyunchimeg Shagdar.

The automated cooperative driving applications require efforts on multiple research domains including robotics, artificial intelligence, and communications to build a safe and intelligent collective driving behavior. While some studies show the potentials of the V2V communications for safer and smoother automated driving, it is still not clear if the standardized technologies can meet the strict requirements of the automated driving applications. More importantly, if the decisions for individual vehicles' control are based on the V2V communications, the communications performance must largely affect the "quality" of the collective behavior. Motivated by this, we study the inter-dependencies between communications and collective automated driving behavior. In our study [43] we combine different V2V communication modes with different dynamic path-finding heuristics, over a population of several hundreds of intelligent vehicles, to observe convergence towards stable traffic. The various traffic stability levels are compared in order to exhibit most efficient combinations of communication modes and path-finding heuristics.

6.15. Distributed Agreement and String Control in Intelligent Vehicular Networks (IVNs)

Participant: Gérard Le Lann.

IVNs are composed of automated (autonomous and communicating) vehicles, ranging from pre-planned platoons to ad hoc vehicular networks (VANETs). Agreement problems in the presence of concurrency and failures are not well investigated yet in IVNs. We have examined a specific class of such problems, those arising in string formations. Regarding string membership (vehicles leaving or joining a string), with few exceptions, safety issues have been addressed so far assuming that (1) no more than 1 insertion operation would be performed at any given time or, (2) every vehicle decides unilaterally, i.e. undertakes a maneuver after having activated some signal, leaving to surrounding vehicles the responsibility of inferring intended maneuvers. Assumption (1) is not realistic. There are numerous risk-prone scenarios where a posteriori reactive approaches (assumption (2)) may fail. Therefore the need for investigating proactive approaches, where vehicles (1) are made aware of intended impending maneuvers, (2) agree on which maneuvers can be safely undertaken, prior to performing physical maneuvers. It follows that a solution to numerous string control problems consists of a pair (A,Φ), where A stands for a distributed agreement algorithm which achieves global coordination in the presence of failures and concurrency, and Φ stands for control laws drawn from control theory and robotics. Algorithm A is run prior to local activations of Φ.
Our work is based on the cohort construct, which serves to formalize the concept of strings. Velocity Agreement is the generic problem selected. At any time, some number of string/cohort members may propose each a new velocity value. In fine, every vehicle computes a unique new velocity V. Proposed values are propagated via neighbor-to-neighbor (N2N) radio communications. We have devised a solution called the VAgree algorithm. In the presence of up to f failures (on-board systems, N2N message losses), the following properties shall hold:

- Validity: Decision value $V = \Psi$(proposed values).
- Agreement: No two members decide differently.
- Time-Bounded Termination: VAgree terminates at most θ time units.
- Synchronicity: Times at which V is posted to on-board systems are comprised within a small time interval ϵ. Distance traveled during ϵ by the member earliest to post V until the latest member does so is an order of magnitude smaller than vehicle sizes.

The VAgree algorithm is presented in a paper which is under submission.

6.16. Standardization and automated vehicles

Participant: Michel Parent.

Michel Parent has been active over the last 4 years in this group to introduce automated vehicles in the scope of service robots and he contributed actively in the activities of several working groups (WGs). In WG7, he participated in the writing of the document ISO13482 on the safety of service robots. This document has reached the Final Draft for an International Standard level (FDIS) and is now published in English and French. It has already been used by companies to certify some robotics products, including Robosoft in France for automated vehicles. In WG8, Michel Parent participated in the elaboration of standard procedures for the testing of service robots and in particular for automated vehicles. The document CD18646 « Robots and robotic devices — Performance criteria and related test methods for service robots — Part 1: Locomotion for wheeled robots » is in progress.

6.17. Legislation and certification of fully automated road vehicles for urban public transport

Participant: Michel Parent.

An important research area of automated road vehicles and one of the focuses of the CityMobil2 Project is to look at the legislation and certification of fully automated road vehicles for urban public transport (the cybercars). This part of the research was done essentially by Michel Parent in 2014 and gave birth to several CityMobil2 deliverables.

One of the tasks was to identify the current legislation in France and the organizations involved in the changes for this legislation. Several meetings were therefore organized at the French level with key persons from the Ministry of Transport, the Ministry of Interior (responsible for the road legislation) and their services (in particular the SRMTG in charge of certifying the guided transport systems). These meetings were essential in obtaining the authorization to operate the cybercars for the demonstration in La Rochelle. At the European level, a meeting was organized in May 2014 with representatives of 12 of the European countries (mostly those involved directly with CityMobil2 or with automated vehicles R&D).

Another task was to propose a certification methodology for automated road transport systems. For this task, a careful analysis of the test site in La Rochelle was conducted and led to a number of use cases. Key elements were defined to perform the risk analysis. Many hazards were identified but the most important ones are the behavior of pedestrian and cyclists. For the analysis, two key variables were defined: the minimum mobile object detection distance (MMODD) and the maximum mobile object speed (MMOS). For each use case, a combination of these 2 variable lead to a maximum vehicle speed in order to reach an acceptable risk evaluated as a combination of severity and probability.

In order to verify the proper behavior of the vehicle itself (lane tracking, obstacle avoidance, comfort,...), a number of standard tests have also been defined and are now proposed at the International level (ISO standards).

6.18. Belief propagation inference for traffic prediction

Participants: Cyril Furtlehner, Jean-Marc Lasgouttes.

This work [60] deals with real-time prediction of traffic conditions in a setting where the only available information is floating car data (FCD) sent by probe vehicles. The main focus is on finding a good way to encode some coarse information (typically whether traffic on a segment is fluid or congested), and to decode it in the form of real-time traffic reconstruction and prediction. Our approach relies in particular on the belief propagation algorithm.

These studies have been done in particular in the framework of the projects Travesti and Pumas.

This year, the work about the theoretical aspects of encoding real valued variables into a binary Ising model has been under review for a Journal and has been largely revised in the process.

6.19. Sparse covariance inverse estimate for Gaussian Markov Random Field

Participants: Cyril Furtlehner, Jean-Marc Lasgouttes.

We investigate the problem of Gaussian Markov random field selection under a non-analytic constraint: the estimated models must be compatible with a fast inference algorithm, namely the Gaussian belief propagation algorithm. To address this question, we introduce the ☆-IPS framework, based on iterative proportional scaling, which incrementally selects candidate links in a greedy manner. Besides its intrinsic sparsity-inducing ability, this algorithm is flexible enough to incorporate various spectral constraints, like e.g. walk summability, and topological constraints, like short loops avoidance. Experimental tests on various datasets, including traffic data from San Francisco Bay Area, indicate that this approach can deliver, with reasonable computational cost, a broad range of efficient inference models, which are not accessible through penalization with traditional sparsity-inducing norms.

This work has been presented at ECML/PKDD 2014 [40]. The code for ☆-IPS has been made available at https://who.rocq.inria.fr/Jean-Marc.Lasgouttes/star-ips/.

6.20. 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* [20].

6.21. Properties of random walks in orthants

Participant: Guy Fayolle.

We pursued works initiated these last years in several directions.

6.21.1. Explicit criterion for the finiteness of the group in the quarter plane

In the book [3], original methods were proposed to determine the invariant measure of random walks in the quarter plane with small jumps, the general solution being obtained via reduction to boundary value problems. Among other things, an important quantity, the so-called *group of the walk*, allows to deduce theoretical features about the nature of the solutions. In particular, when the *order* of the group is finite, necessary and sufficient conditions have been given in [3] for the solution to be rational or algebraic. When the underlying algebraic curve is of genus 1, we propose, in collaboration with R. Iasnogorodski (St-Petersbourg, Russia), a concrete criterion ensuring the finiteness of the group. It turns out that this criterion is always tantamount to the cancellation of a single constant, which can be expressed as the determinant of a matrix of order 3 or 4, and depends in a polynomial way on the coefficients of the walk [55].

6.21.2. About a possible analytic approach for walks with arbitrary big jumps in $\mathbb{Z}_+^2$

The article [21], achieved in collaboration with K. Raschel (CNRS and University F. Rabelais, Tours) considers random walks with arbitrary big jumps. For that class of models, we announce a possible extension of the analytic approach proposed in [3], initially valid for walks with small steps in the quarter plane. New technical challenges arise, most of them being tackled in the framework of generalized boundary value problems on compact Riemann surfaces.

6.21.3. Correction of papers

Guy Fayolle found important errors in several articles dealing with models involving random walks in $\mathbb{Z}_+^2$. This is the object of the letter to the editors [19]. The concerned authors have provided new correct versions of their studies.

6.21.4. Communication networks with harvesting energy supply

In collaboration with S. Foss (Heriot-Watt University, Edinburgh), we started to analyze stability and performance of a number of models of parallel queues with multiple access and individual energy supplies. Energy limitation in general decreases the stability region, but also may increase it for specific parameter regions. The most difficult and intriguing cases arise when the input rates of requests and of energy items are close. Preliminary models of physical interest involve random walks in $\mathbb{Z}_+^4$.

6.22. Global optimization for online resource allocation

Participant: Jean-Marc Lasgouttes.

As part of the Mobility 2.0 FP7 project, we have considered the possibility to allocate charging stations to Full Electric Vehicle (FEV) users in a way that, instead of merely minimizing their travel time, tries to improve the travel time for the whole community.

The aim of the global optimization algorithm is to pursue the minimization of the mean squared travel time encountered by each user. Our setting can be seen as a resource allocation problem, known as the "Transportation Problem" in Operational Research literature. It is solvable using several algorithms, among which the simplex algorithm. Unfortunately, these algorithms are not usable here for two reasons:

- The allocation of slots to the users is done online, when the user does a request. It is not possible to wait until all the users are known before doing the allocation;
- The complexity of these algorithms is very high, especially since, due to the effect of range limitations, each request has different characteristics.

We therefore present here a simplified approach, which should be fast enough to scale for large systems. The principle of the algorithm is to penalize the cost for the user with an approximation of the extra cost incurred to future users who compete for the same resource (a charging or parking slot). Since the implications can be intricate, we only consider a first order effect.

Our work in the Mobility 2.0 project has been briefly described in [37].

SMIS Project-Team

6. New Results

6.1. Flash-Based Data Management

Participants: Nicolas Anciaux, Matias Bjørling, Philippe Bonnet, Luc Bouganim [correspondent], Niv Dayan, Saliha Lallali, Philippe Pucheral, Iulian Sandu Popa.

There is a long tradition of work around the understanding and optimization of NAND Flash memory in the team (e.g., [7], [9]). Current work in this area covers the optimization of SSD use in DBMS engines and the design of Flash-based indexing techniques for textual and spatio-temporal data. These works on Flash-Based indexing complete the work initiated in the last years on the storage and indexing engine of PlugDB (not repeated in this report but the interested reader is referred to a DAPD'14 journal publication detailing these techniques [14]).

Flash storage optimization. Solid State Drives (SSDs), based on flash chips, are now the secondary storage of choice for data intensive applications. Database systems can now rely on high performance SSDs to store log, indexes and data either on servers or in the cloud. While SSDs provide increasingly high performance out of the box, maintaining high throughput and low latency as the utilization of SSDs increases and despite abrupt changes in the workload remains a challenge. This question is central for database designers and administrators, cloud service providers, and SSD constructors. The answer depends on write-amplification, i.e., garbage collection overhead. More specifically, the answer depends on how write-amplification evolves in time. We derived a mathematical expression that relates over provisioning to write-amplification. We introduced a new block manager, called Wolf, or WOrkload Leveler for Flash. Wolf is able to detect and quickly adapt to changes in workload by pro-actively reallocating over-provisioned space among the groups based on their changing needs. It adapts better to stable workloads by measuring the update frequencies of groups instead of making assumptions about them. It uses a novel near-optimal closed-form expression to allocate over-provisioned space to groups.

Flash-based keyword indexing. As smart objects gain the capacity to acquire, store and process large volumes of data, new services emerge. However, the smart objects have to be endowed with typical data management capabilities to enable all these services. In this work, we revisit the traditional problem of information retrieval queries over large collections of files in an embedded context. A file can be any form of document, picture or data stream associated with a set of terms. A query can be any form of keyword search using a ranking function (e.g., TF-IDF) identifying the top-k most relevant files. The proposed search engine can be used in sensors to search for relevant objects in their surroundings, in cameras to search pictures by using tags, in personal smart dongles to secure the querying of documents and files hosted in an untrusted Cloud or in smart meters to perform analytic tasks (i.e., top-k queries) over sets of events (i.e., terms) captured during time windows (i.e., files) [21]. Designing such embedded search engine is however challenging due to a combination of severe and conflicting hardware constraints (e.g., a tiny RAM combined with a NAND Flash persistent storage badly adapted to random fine-grain updates). To tackle this challenge, we introduce three design principles, namely Write-Once Partitioning, Linear Pipelining and Background Linear Merging, and show how they can be combined to produce an embedded search engine reconciling high insert/delete/update rate and query scalability. We have implemented our search engine on a development board having a hardware configuration representative for smart objects. The experimental results demonstrate the scalability of the approach and its superiority compared to state of the art methods [28]. This work is part of Saliha Lallali's Ph.D. thesis.

Flash-based spatio-temporal indexing. The convergence of mobile computing, wireless communications and sensors has raised the development of many applications exploiting a massive flow of spatio-temporal data such as location-based services, participatory sensing, or traffic management [15]. Among the most active research topics in this area is the spatio-temporal data indexing. Nevertheless, since a few years a new fundamental parameter has made its entry on the database scene: the NAND flash storage. However, the peculiar characteristics of flash memory require redesigning the existing data storage and indexing techniques that were devised for magnetic hard-disks. In this study we propose TRIFL, an efficient and generic TRajectory Index for FLash. TRIFL is designed around the key requirements of trajectory indexing and flash storage. TRIFL is generic in the sense that it is efficient for both simple flash storage devices such as the SD cards and more powerful devices such as the solid state drives. In addition, TRIFL is supplied with an online self-tuning algorithm that allows adapting the index structure to the workload and the technical specifications of the flash storage device to maximize the index performance. Moreover, TRIFL achieves good performance with relatively low memory requirements, which makes the index appropriate for many application scenarios. The experimental evaluation shows that TRIFL outperforms the representative indexing methods on magnetic disks and flash disks. This work is part of Dai-Hai Ton That Ph.D. thesis, co-supervised by Iulian Sandu Popa.

6.2. Secure Global Computing on Asymmetric Architecture

Participants: Benjamin Nguyen [correspondent], Philippe Pucheral, Quoc-Cuong To.

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 studied two different computing paradigms on this architecture: our first contribution was to study the execution of Privacy Preserving Data Publishing (PPDP) algorithms on such an architecture, and provided generic protocols to deal with all kinds of PPDP algorithms, which are robust against honest-but-curious and malicious adversaries [12], including vulgarization aspects [25]. Our second contribution was to study general SQL queries in this same execution context. For now, we have concentrated on the subset of SQL queries without joins, but including Group By and aggregates, and show how to secure their execution in the presence of honest-but-curious attackers [19]. Cost models and experiments demonstrate that this approach can scale to nationwide infrastructures [20][16]. This work is part of Quoc-Cuong To's Ph.D. thesis started in sept. 2012, and should be extended in particular to cover joins. We also plan to extend this general framework through a collaboration with INSA Centre Val de Loire, LIFO Lab and University of Paris Nord, LIPN lab, to study the secure execution of Map/Reduce on the Asymmetric Architecture.

6.3. Personal Cloud

Participants: Nicolas Anciaux [correspondent], Luc Bouganim, Athanasia Katsouraki, Benjamin Nguyen, Philippe Pucheral, Iulian Sandu Popa, Paul Tran Van.

We are witnessing an exponential increase in the acquisition of personal data about the individuals or produced by them. Today, this information is managed using Web applications, centralizing this data in cloud data servers, under the control of few Web majors [5]. However, it has now become clear that (1) centralizing millions of personal records exposes the data to very sophisticated attacks, linked to a very high potential benefit in case of success (millions of records being revealed), and (2) delegating the management of personal records without any tangible guarantee for the individuals leads to privacy violations, the data being potentially

made accessible to other organizations (e.g., governments, commercial partners) and being subject to lucrative secondary usages (not advertised to the individuals). To face this situation, many recent initiatives push towards the emergence of the Personal Cloud paradigm. A personal cloud can be viewed as a personal server, owned by a given individual, which gives to its owner the ability to store her complete digital environment, synchronize it among various devices and share it with other individuals and applications under control. Many projects and startups currently investigate this solution, like OpenPDS, CozyCloud, OwnCloud, etc. In the SMIS team, we claim the need of a Secure Personal Cloud, and promote the introduction of a secure (tamper resistant) data engine in the architecture [11]. On this basis, we investigate new data sharing and dissemination models, where usage and access control rules endorsed by the individuals could be enforced. In 2014, we have presented this vision at EDBT'14 [18]. Several underlying research problems and perspectives have been presented in [11]. We have started a cooperation with the startup CozyCloud at the end of 2014. A contract was signed at the end of 2014 to integrate PlugDB in a CozyCloud instance and the PhD of Paul Tran Van (CIFRE SMIS-CozyCloud) has just started to explore new data sharing techniques which could be enforced in the secure personal cloud model. Athanasia Katsouraki is working on privacy issues and on adoption of the secure data engine in cooperation with the economists (CERDI) in the context of the Digital Society Institute (DSI).

6.4. Folk-IS

Participants: Nicolas Anciaux [correspondent], Luc Bouganim, Philippe Pucheral.

According to many studies, IT should become a key facilitator in establishing primary education, reducing mortality or supporting commercial initiatives in Least Developed Countries. 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, as well as political commitment. In [5], we proposed the vision of trusted cells, a data platform for personal data services where the shared infrastructure (typically the cloud) is untrusted, while personal devices (such as smart phones, tablets or set-top box) are trusted execution environments. We revisited this vision to the context of LDCs. We proposed a new paradigm, that we call Folk-enabled Information System (Folk-IS), based on a fully decentralized and participatory approach, where each individual implements a small subset of a complete information system without the need for a shared networked infrastructure. As trusted cells, Folk-IS builds upon the emergence of highly secure, portable and low-cost storage and computing devices, called hereafter Smart Tokens. Here however, the focus is on low-cost of ownership, deployment and maintenance, and on the absence of a networked infrastructure. With Folk-IS and thanks to their smart tokens, people will transparently and opportunistically perform data management and networking tasks as they physically move, so that IT services are truly delivered by the crowd. We have published in [17] the Folk-IS vision and main principles, and in [13] a more detailed paper including technical challenges, specific to that approach and an exploitation and feasibility analysis of the Folk-IS vision.

WILLOW Project-Team

6. New Results

6.1. Highlights of the Year

- J. Sivic started ERC project LEAP (2014-2018).
- J. Sivic serves as a Program Chair for International Conference on Computer Vision, Santiago, Chile, 2015

6.2. 3D object and scene modeling, analysis, and retrieval

(a) Internet paintings and drawings. (b) Painting viewpoints. (c) Aligned painting to 3D model.

Figure 1. Our system automatically aligns and recovers the viewpoint of paintings, drawings, and historical photographs to a 3D model of an architectural site.

6.2.1. 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 published at ACM Transactions on Graphics 2014 [3] and its extension has appeared at RFIA 2014 [17]. The problem addressed in this work is illustrated in Figure 1 and example results are shown in Figure 2 .

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.

6.2.2. 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.

This work poses object category detection in images as a type of 2D-to-3D alignment problem, utilizing the large quantities of 3D CAD models that have been made publicly available online. Using the "chair" class as a running example, we propose an exemplar-based 3D category representation, which can explicitly model chairs of different styles as well as the large variation in viewpoint. We develop an approach to establish part-based correspondences between 3D CAD models and real photographs. This is achieved by (i) representing each 3D model using a set of view-dependent mid-level visual elements learned from synthesized views in a discriminative fashion, (ii) carefully calibrating the individual element detectors on a common dataset of negative images, and (iii) matching visual elements to the test image allowing for small mutual deformations but preserving the viewpoint and style constraints. We demonstrate the ability of our system to align 3D models with 2D objects in the challenging PASCAL VOC images, which depict a wide variety of chairs in complex scenes. This work has been published at CVPR 2014 [9].

6.2.3. Anisotropic Laplace-Beltrami Operators for Shape Analysis

Participants: Mathieu Andreux [TUM], Emanuele Rodola [TUM], Mathieu Aubry, Daniel Cremers [TUM].

This work introduces an anisotropic Laplace-Beltrami operator for shape analysis. While keeping useful properties of the standard Laplace-Beltrami operator, it introduces variability in the directions of principal curvature, giving rise to a more intuitive and semantically meaningful diffusion process. Although the benefits of anisotropic diffusion have already been noted in the area of mesh processing (e.g. surface regularization), focusing on the Laplacian itself, rather than on the diffusion process it induces, opens the possibility to effectively replace the omnipresent Laplace-Beltrami operator in many shape analysis methods. After providing a mathematical formulation and analysis of this new operator, we derive a practical implementation on discrete meshes. Further, we demonstrate the effectiveness of our new operator when employed in conjunction with different methods for shape segmentation and matching. This work has been published at the Sixth Workshop on Non-Rigid Shape Analysis and Deformable Image Alignment (NORDIA) 2014 [8].

6.2.4. 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 published at CVPR 2014 [15].

6.2.5. On Image Contours of Projective Shapes

Participants: Jean Ponce, Martial Hebert [CMU].

This work revisits classical properties of the outlines of solid shapes bounded by smooth surfaces, and shows that they can be established in a purely projective setting, without appealing to Euclidean measurements such as normals or curvatures. In particular, we give new synthetic proofs of Koenderink's famous theorem on convexities and concavities of the image contour, and of the fact that the rim turns in the same direction as the viewpoint in the tangent plane at a convex point, and in the opposite direction at a hyperbolic point. This suggests that projective geometry should not be viewed merely as an analytical device for linearizing calculations (its main role in structure from motion), but as the proper framework for studying the relation between solid shape and its perspective projections. Unlike previous work in this area, the proposed approach does not require an oriented setting, nor does it rely on any choice of coordinate system or analytical considerations. This work has been published at ECCV 2014 [14].

6.3. Category-level object and scene recognition

6.3.1. Finding Matches in a Haystack: A Max-Pooling Strategy for Graph Matching in the Presence of Outliers

Participants: Minsu Cho, Jian Sun, Olivier Duchenne, Jean Ponce.

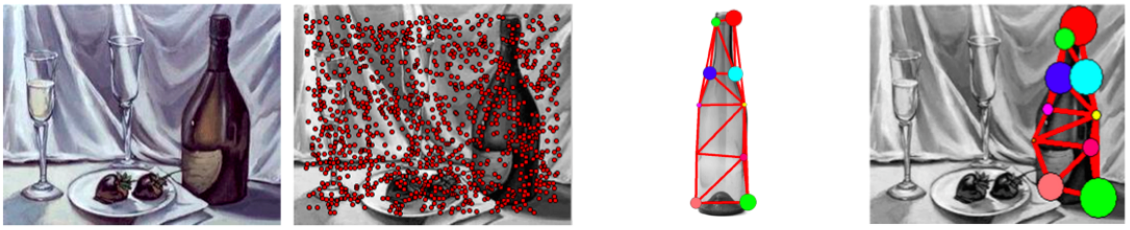


Figure 3. Feature matching in the presence of outliers. (a) In real-world scenes, background clutter often produces numerous outlier features, making it hard to find correspondences. (b) We address the issue with a max-pooling approach to graph matching. The proposed method is not only resilient to deformations but also remarkably tolerant to outliers. Each node on the left image corresponds to one with the same color on the right image, where bigger nodes represent more similar nodes. (Best viewed in color.)

A major challenge in real-world feature 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 harder to distinguish inliers from outliers due to clutters. In this work, we propose a max-pooling approach to graph matching, which is not only resilient to deformations but also remarkably tolerant to outliers. 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 published at CVPR 2014 [11]. The proposed method and its qualitative results are illustrated in Figure 3 .

6.3.2. *Unsupervised Object Discovery and Localization in the Wild: Part-based Matching with Bottom-up Region Proposals*

Participants: Minsu Cho, Suha Kwak, Cordelia Schmid [Inria Lear], Jean Ponce.

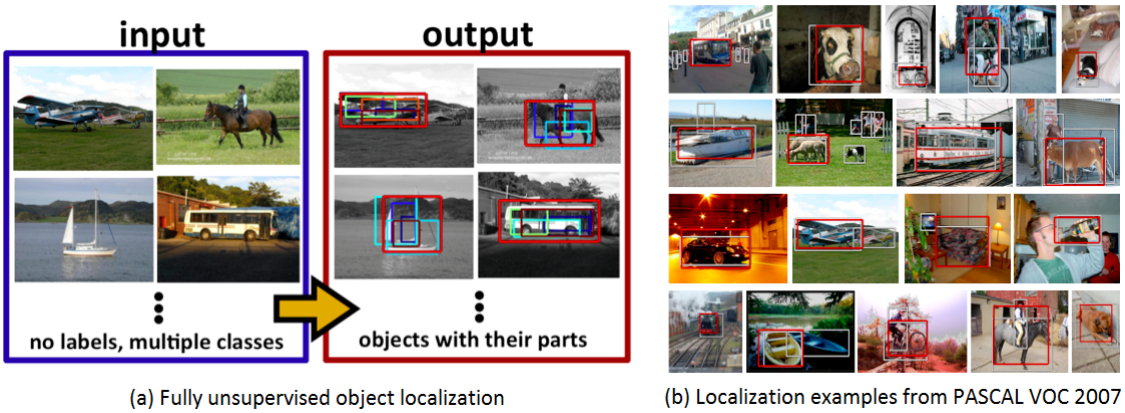


Figure 4. Unsupervised object discovery in the wild. (a) We tackle object localization in an unsupervised scenario without any types of annotations, where a given image collection may contain multiple dominant object classes and even outlier images. The proposed method discovers object instances (red bounding boxes) with their distinctive parts (smaller boxes). (b) Examples of localization on mixed-class PASCAL VOC 2007 train/val datasets are shown. Red boxes represent localized objects while white boxes are ground truth annotations. (Best viewed in color.)

This work addresses unsupervised discovery and localization of dominant objects from a noisy image collection of multiple object classes. The setting of this problem is fully unsupervised, without even image-level annotations or any assumption of a single dominant class. This is significantly more general than typical colocalization, cosegmentation, or weakly-supervised localization tasks. We tackle the discovery and localization problem using a part-based matching approach: We use off-the-shelf region proposals to form a set of candidate bounding boxes for objects and object parts. These regions are efficiently matched across images using a probabilistic Hough transform that evaluates the confidence in each candidate region considering both appearance similarity and spatial consistency. Dominant objects are discovered and localized by comparing the scores of candidate regions and selecting those that stand out over other regions containing them. Extensive experimental evaluations on standard benchmarks demonstrate that the proposed approach significantly outperforms the current state of the art in colocalization, and achieves robust object discovery in

challenging mixed-class datasets. This work has been submitted to CVPR 2015 [22]. The proposed method and its qualitative results are illustrated in Figure 4 .

6.3.3. 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. This work has been published at CVPR 2014 [13].

6.3.4. Weakly supervised object recognition with convolutional neural networks

Participants: Maxime Oquab, Leon Bottou [MSR New York], Ivan Laptev, Josef Sivic.

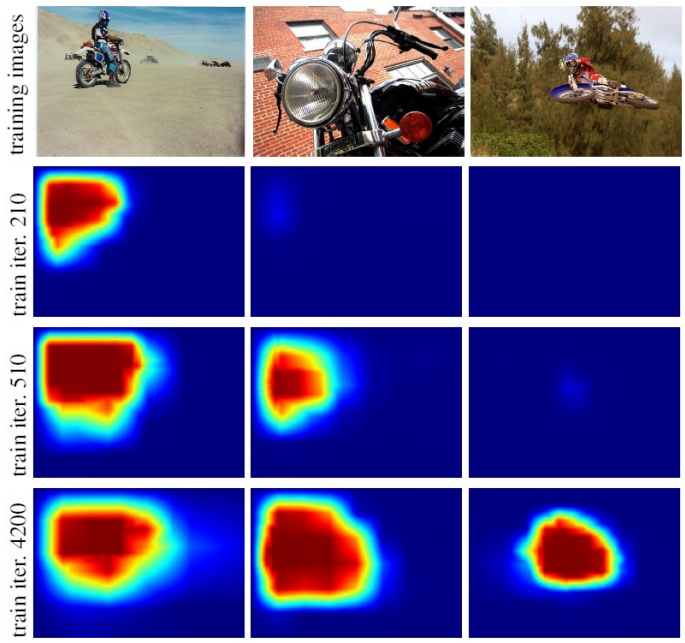


Figure 5. Evolution of localization score maps for the motorbike class over iterations of our weakly-supervised CNN training. Note that locations of objects with more usual appearance are discovered earlier during training.

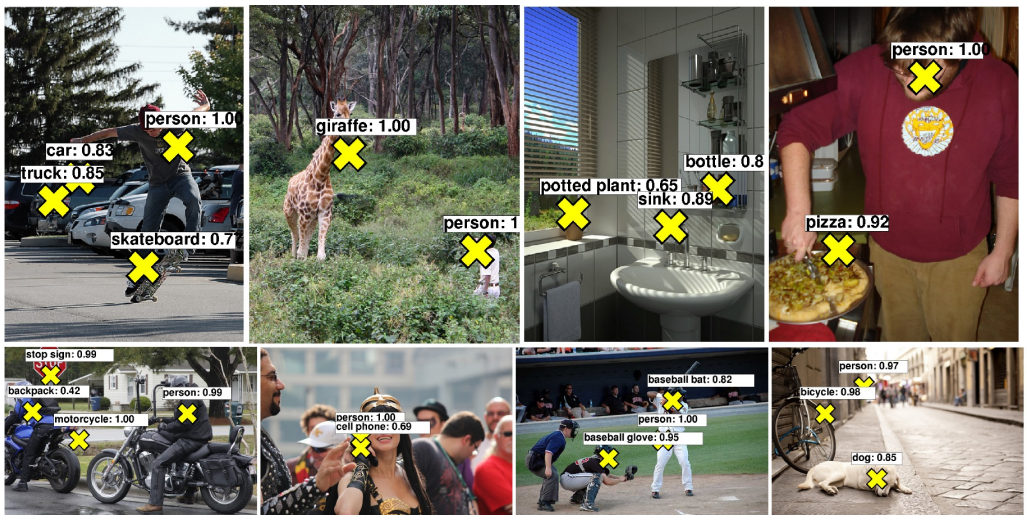


Figure 6. Example location predictions for images from the Microsoft COCO validation set obtained by our weakly-supervised method. Note that our method does not use object locations at training time, yet can predict locations of objects in test images (yellow crosses). The method outputs the most confident location for most confident object classes.

Successful methods for visual object recognition typically rely on training datasets containing lots of richly annotated images. Detailed image annotation, e.g. by object bounding boxes, however, is both expensive and often subjective. We describe a weakly supervised convolutional neural network (CNN) for object classification that relies only on image-level labels, yet can learn from cluttered scenes containing multiple objects (see Figure 5). We quantify its object classification and object location prediction performance on the Pascal VOC 2012 (20 object classes) and the much larger Microsoft COCO (80 object classes) datasets. We find that the network (i) outputs accurate image-level labels, (ii) predicts approximate locations (but not extents) of objects, and (iii) performs comparably to its fully-supervised counterparts using object bounding box annotation for training. This work has been submitted to CVPR 2015 [23]. Illustration of localization results by our method in Microsoft COCO dataset is illustrated in Figure 6 .

6.3.5. *Learning Dictionary of Discriminative Part Detectors for Image Categorization and Cosegmentation*

Participants: Jian Sun, Jean Ponce.

This work proposes a novel approach to learning mid-level image models for image categorization and cosegmentation. We represent each image class by a dictionary of discriminative part detectors that best discriminate that class from the background. We learn category-specific part detectors in a weakly supervised setting in which the training images are only labeled with category labels without part / object location labels. We use a latent SVM model regularized by l1,2 group sparsity to learn the discriminative 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 learned part detectors to image classification and cosegmentation, and quantitative experiments with standard benchmarks show that our approach matches or improves upon the state of the art. This work has been submitted to PAMI [24].

6.4. Image restoration, manipulation and enhancement

6.4.1. Fast Local Laplacian Filters: Theory and Applications

Participants: Mathieu Aubry, Sylvain Paris [Adobe], Samuel Hasinoff [Google], Jan Kautz [University College London], Fredo Durand [MIT].

Multi-scale manipulations are central to image editing but they are also prone to halos. Achieving artifact-free results requires sophisticated edge-aware techniques and careful parameter tuning. These shortcomings were recently addressed by the local Laplacian filters, which can achieve a broad range of effects using standard Laplacian pyramids. However, these filters are slow to evaluate and their relationship to other approaches is unclear. In this work, we show that they are closely related to anisotropic diffusion and to bilateral filtering. Our study also leads to a variant of the bilateral filter that produces cleaner edges while retaining its speed. Building upon this result, we describe an acceleration scheme for local Laplacian filters on gray-scale images that yields speed-ups on the order of 50x. Finally, we demonstrate how to use local Laplacian filters to alter the distribution of gradients in an image. We illustrate this property with a robust algorithm for photographic style transfer. This work has been published at ACM Transactions on Graphics 2014 [2].

6.4.2. Learning a Convolutional Neural Network for Non-uniform Motion Blur Removal

Participants: Jian Sun, Wenfei Cao, Zongben Xu, Jean Ponce.

In work work, we address the problem of estimating and removing non-uniform motion blur from a single blurry image. We propose a deep learning approach to predicting the probabilistic distribution of motion blur at the patch level using a convolutional neural network (CNN). We further extend the candidate set of motion kernels predicted by the CNN using carefully designed image rotations. A Markov random field model is then used to infer a dense non-uniform motion blur field enforcing the motion smoothness. Finally the motion blur is removed by a non-uniform deblurring model using patch-level image prior. Experimental evaluations show that our approach can effectively estimate and remove complex non-uniform motion blur that cannot be well achieved by the previous approaches. This work has been submitted to CVPR 2015.

6.5. Human activity capture and classification

6.5.1. Weakly Supervised Action Labeling in Videos Under Ordering Constraints

Participants: Piotr Bojanowski, Remi Lajugie [Inria Sierra], Francis Bach [Inria Sierra], Ivan Laptev, Jean Ponce, Cordelia Schmid [Inria Lear], Josef Sivic.

We are given a set of video clips, each one annotated with an ordered list of actions, such as “walk” then “sit” then “answer phone” extracted from, for example, the associated text script. We seek to temporally localize the individual actions in each clip as well as to learn a discriminative classifier for each action. We formulate the problem as a weakly supervised temporal assignment with ordering constraints. Each video clip is divided into small time intervals and each time interval of each video clip is assigned one action label, while respecting the order in which the action labels appear in the given annotations. We show that the action label assignment can be determined together with learning a classifier for each action in a discriminative manner. We evaluate the proposed model on a new and challenging dataset of 937 video clips with a total of 787720 frames containing sequences of 16 different actions from 69 Hollywood movies. This work has been published at ECCV 2014 [10].

6.5.2. Predicting Actions from Static Scenes

Participants: Tuan-Hung Vu, Catherine Olsson [MIT], Ivan Laptev, Aude Oliva [MIT], Josef Sivic.

Human actions naturally co-occur with scenes. In this work we aim to discover action-scene correlation for a large number of scene categories and to use such correlation for action prediction. Towards this goal, we collect a new SUN Action dataset with manual annotations of typical human actions for 397 scenes. We next discover action-scene associations and demonstrate that scene categories can be well identified from their associated actions. Using discovered associations, we address a new task of predicting human actions for images of static scenes. We evaluate prediction of 23 and 38 action classes for images of indoor and outdoor scenes respectively and show promising results, see Figure 7 . We also propose a new application of geo-localized action prediction and demonstrate ability of our method to automatically answer queries such as “Where is a good place for a picnic?” or “Can I cycle along this path?”. This work has been published in ECCV 2014 [16].

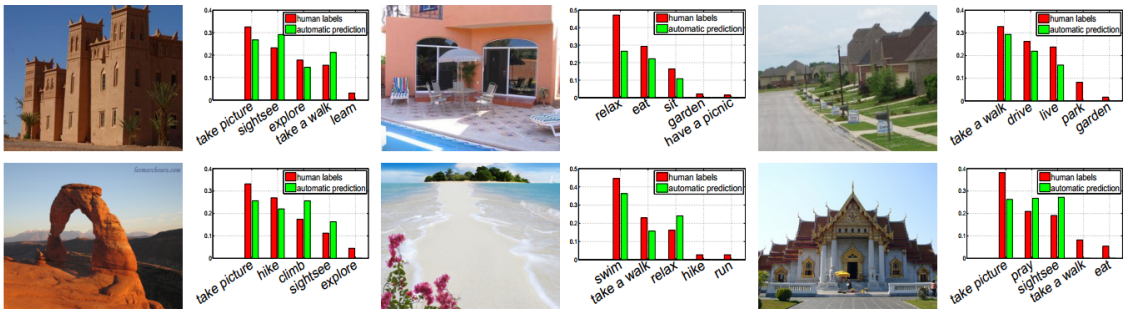


Figure 7. Automatic visual action prediction for test images in SUN Action dataset.

6.5.3. *Efficient feature extraction, encoding and classification for action recognition*

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 continuously improved over the recent years, the low speed of feature extraction and subsequent recognition prevents current methods from scaling up to real-size problems. We address this issue and first develop highly efficient video features using motion information in video compression. We next explore feature encoding by Fisher vectors and demonstrate accurate action recognition using fast linear classifiers. Our method improves the speed of video feature extraction, feature encoding and action classification by two orders of magnitude at the cost of minor reduction in recognition accuracy. We validate our approach and compare it to the state of the art on four recent action recognition datasets. This work has been published at CVPR 2014 [12].

6.5.4. *On Pairwise Cost for Multi-Object Network Flow Tracking*

Participants: Visesh Chari, Simon Lacoste-Julien [Inria Sierra], Ivan Laptev, Josef Sivic.

Multi-object tracking has been recently approached with the min-cost network flow optimization techniques. Such methods simultaneously resolve multiple object tracks in a video and enable modeling of dependencies among tracks. Min-cost network flow methods also fit well within the “tracking-by-detection” paradigm where object trajectories are obtained by connecting per-frame outputs of an object detector. Object detectors, however, often fail due to occlusions and clutter in the video. To cope with such situations, we propose an approach that regularizes the tracker by adding second order costs to the min-cost network flow framework. While solving such a problem with integer variables is NP-hard, we present a convex relaxation with an efficient rounding heuristic which empirically gives certificates of small suboptimality. Results are shown on real world video sequences and demonstrate that the new constraints help selecting longer and more accurate tracks improving over the baseline tracking-by-detection method. This work has been submitted to CVPR 2015 [21].