Activity Report 2011

Section New Results
ALGORITHMS, PROGRAMMING, SOFTWARE AND ARCHITECTURE
1. ABSTRACTION Project-Team .............................................................. 5
2. ALGORITHMS Project-Team ............................................................ 11
3. AOSTE Project-Team ...................................................................... 12
4. CASCADE Project-Team (section vide) ............................................. 17
5. CONTRAINTES Project-Team ........................................................... 18
6. FORMES Team ............................................................................. 22
7. GALLIUM Project-Team ................................................................. 29
8. MOSCOVA Project-Team ................................................................. 36
9. PARKAS Team ............................................................................. 41
10. PLR2 Project-Team ...................................................................... 45
11. SALSA Project-Team .................................................................... 51
12. SECRET Project-Team ................................................................... 55
13. CAD Team ................................................................................... 59
14. CLASSIC Project-Team ................................................................. 76
15. GAMMA3 Project-Team ................................................................. 79
16. MATHFI Project-Team ................................................................. 84
17. MICMAC Project-Team ................................................................. 88
18. POEMS Project-Team .................................................................. 96
19. SIERRA Project-Team .................................................................. 114

APPLIED MATHEMATICS, COMPUTATION AND SIMULATION
20. BANG Project-Team .................................................................... 121
21. CLIME Project-Team .................................................................. 133
22. ESTIME Project-Team (section vide) ........................................... 147
23. MACS Project-Team ................................................................... 148
24. NEUROMATHCOMP Project-Team .............................................. 151
25. REO Project-Team .................................................................... 158
26. SISYPHE Project-Team ............................................................... 162

COMPUTATIONAL SCIENCES FOR BIOLOGY, MEDICINE AND THE ENVIRONMENT
27. ARLES Project-Team .................................................................... 167
28. GANG Project-Team .................................................................... 176
29. HIPERCOM Project-Team ............................................................ 184
30. RAP Project-Team ..................................................................... 194
31. REGAL Project-Team .................................................................. 199
32. TREC Project-Team ................................................................... 207

NETWORKS, SYSTEMS AND SERVICES, DISTRIBUTED COMPUTING
33. ALPAGE Project-Team ................................................................. 221
34. AXIS Project-Team ..................................................................... 231
35. IMARA Project-Team .................................................................. 239
36. IMEDIA Project-Team ................................................................. 245

PERCEPTION, COGNITION, INTERACTION
37. SMIS Project-Team ................................................................. 253
38. WILLOW Project-Team ......................................................... 255
6. New Results

6.1. Abstractions of Functions

Participants: Patrick Cousot, Radhia Cousot.

The idea of domain segmentation for arrays [18] has been extended to the abstraction of functions [41] by combination of a partionning of their domain of definition and a functional or relational abstraction of blocks into their co-domain [17].

6.2. Analysis of Biological Pathways

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

6.2.1. Automatic Reduction of Differential Semantics

Participants: Ferdinanda Camporesi, Vincent Danos [University of Edinburgh], Jérôme Feret, Walter Fontana [Harvard Medical School], Russ Harmer [Harvard Medical School], Jean Krivine [Paris VII].

We have developed an abstract interpretation-based framework that enables the reduction of the differential semantics for protein-protein interaction networks. Results are sound since trajectories in the abstract system are projections of the trajectories in the concrete system.

The flow of information is a key element in our model reduction framework because it enables the identification of the correlations which are useless when computing observables of interest. Thus there is a need of providing good trade-off in the description of the flow of information throughout the biochemical structure of chemical species.

The notion of symmetries between sites is also important, since knowing that two sites have exactly the same capabilities of interaction enable exact quotienting (or lumping) of the set of reachable species.

In [13], [14], we have proposed a heterogeneous over-approximation of the flow of information where the flow that is attached to an agent can depend on its relative position in a chemical species. Moreover, we have showed how to use symmetries between sites so as to define another model reduction and we have proposed an algebraic product to combine model reductions, the product of two reduced models being the least abstract model which is at least as abstract as both model.

6.2.2. Automatic Reduction of Stochastic Semantics

Participants: Ferdinanda Camporesi, Jérôme Feret, Thomas Henzinger [Institute of Science and Technology, Austria], Heinz Koeppl [ETH Zürich], Tatjana Petrov [ETH Zürich].

We have proposed an abstract interpretation-based framework for reducing the state-space of stochastic semantics for protein-protein interaction networks. Our framework ensures that the trace distribution of the reduced system is the exact projection of the trace distribution of the concrete system. Moreover, when the abstraction is complete, if any pair of concrete states that have the same abstraction are equiprobable at initial state, any pair of concrete states that share the same abstraction are equiprobable at any time $t$.

In [12], we have formalized the model reduction framework for the stochastic semantics and we have established the relationships with the notions of lumpability, and bisimulation is established.
6.3. Automatic Array Content Analysis by Segmentation

Participants: Patrick Cousot, Radhia Cousot, Francesco Logozzo [Microsoft Research (Redmond, USA)].

In [18], we introduce FunArray, a parametric segmentation abstract domain functor for the fully automatic and scalable analysis of array content properties. The functor enables a natural, painless and efficient lifting of existing abstract domains for scalar variables to the analysis of uniform compound data-structures such as arrays and collections (as well as matrices when instantiating the functor on itself). The analysis automatically and semantically divides arrays into consecutive non-overlapping possibly empty segments. Segments are delimited by sets of bound symbolic expressions and abstracted uniformly. All bound expressions appearing in a set are equal in the concrete. The FunArray can be naturally combined via reduced product with any existing analysis for scalar variables. The bound expressions, the segment abstractions and the reduction operator are the three parameters of the analysis. Once the functor has been instantiated with fixed parameters, the analysis is fully automatic.

We first prototyped FunArray in Arrayal to adjust and experiment with the abstractions and the algorithms to obtain the appropriate precision/ratio cost. Then it was implemented into CCCHECK (formerly CLOUSOT), an abstract interpretation-based static contract checker for .NET by Francesco Logozzo. The precision and the performance of the analysis has been empirically validated by running it on the main libraries of .NET and on its own code. It was able to infer thousands of invariants and to verify the implementation with a modest overhead (circa 1%). To the best of our knowledge this is the first analysis of this kind applied to such a large code base, and proven to scale.

6.4. Extrapolation operators for combinations of abstract domains

Participants: Agostino Cortesi [Università Ca’Foscari di Venizia], Matteo Zanioli.

Extrapolation operators are crucial to ensure the scalability of the analysis to large software systems. In [10], we set the ground for a systematic design of widening and narrowing operators, by comparing the different definitions introduced in the literature and by discussing how to tune them in case of domain abstraction and domains’ combination through Cartesian and reduced products.

6.5. Grammar Semantics, Analysis and Parsing

Participants: Patrick Cousot, Radhia Cousot.

In [11], we study the abstract interpretations of a fixpoint protoderivation semantics defining the maximal derivations of a transitional semantics of context-free grammars akin to pushdown automata. The result is a hierarchy of bottom-up or top-down semantics refining the classical equational and derivational language semantics and including Knuth grammar problems, classical grammar flow analysis algorithms, and parsing algorithms.

6.6. Information Flow

The analysis of the flow of information in a program consists in detecting the propagation of sensitive information through the program points of this program thanks to a dependency analysis.

6.6.1. Dependency Analysis and Numerical Invariants

Participants: Agostino Cortesi [Università Ca’Foscari di Venizia], Matteo Zanioli.

A new framework has been proposed in [16], that combines variable dependency analysis, based on propositional formulas, and variables’ value analysis, based on generic numerical domains.

6.6.2. Leakage Analysis

Participants: Matteo Zanioli [Correspondent], Pietro Ferrara [ETH, Zurich], Agostino Cortesi [Università Ca’ Foscari].
In [24], we present SAILS, a new tool that combines SAMPLE, a generic static analyzer, and a sophisticated domain for leakage analysis. This tool does not require to modify the original language, since it works with mainstream languages like JAVA™, and it does not require any manual annotation. SAILS can combine the information leakage analysis with different heap abstractions, inferring information leakage over programs with complex data structures. SAILS has been applied to the analysis of the SecuriBench-micro suite. The experimental results underline the effectiveness of the analysis, since SAILS is in position to analyze several benchmarks in about 1 second without producing false alarms in more than 90% of the programs.

### 6.7. Linear Absolute Value Relation Analysis

**Participants:** Liqian Chen [National Laboratory for Parallel and Distributed Processing, Changsha, P. R. China], Antoine Miné, Ji Wang [National Laboratory for Parallel and Distributed Processing, Changsha, P. R. China], Patrick Cousot.

We present in [15] an abstract domain dealing with linear inequalities involving variables together with their absolute values. It is an extension of the classical linear relation analysis, which permits to deal with some non convex numerical sets. A first nice result states the equivalence between these “linear absolute value inequalities” (AVI) and “interval linear inequalities”, and “extended linear complementary inequalities” (XLCP, pairs of positive solutions whose pairwise components are not both not zero). The key contribution is the extension of the double-description of polyhedra to XLCP solutions, which is then used to define the standard operations on AVI. The method has been implemented, and experiments show interesting results, with reasonable performances with respect to linear relation analysis.

### 6.8. Probabilistic Analysis

**Participants:** Patrick Cousot, Michaël Monerau.

The abstract interpretation theory has been widely used in the past decades for verifying properties of computer systems. We have introduced a new extension of this well-known framework to the case of probabilistic systems [21].

The probabilistic abstraction framework we propose allows to systematically lift any classical analysis or verification method to the probabilistic setting by separating in the program semantics the probabilistic behavior from the (non-)deterministic behavior. This separation provides new insights for designing novel probabilistic static analyses and verification methods.

We have defined concrete probabilistic semantics and proposed different ways to abstract them. The approach is expressive and effective. The previous techniques for probabilistic analysis are actually abstractions expressible in our framework.

### 6.9. Safety

**Participants:** Patrick Cousot, Radhia Cousot.

The abstract interpretation design principle has been applied to the design of new forward and backward proof, verification and analysis methods for safety [17]. The safety collecting semantics defining the strongest safety property of programs is first expressed in a constructive fixpoint form. Safety proof and checking/verification methods then immediately follow by fixpoint induction. Static analysis of abstract safety properties such as invariance are constructively designed by fixpoint abstraction (or approximation) to (automatically) infer safety properties.

### 6.10. Security

**Participants:** Patrick Cousot, Radhia Cousot.
We have developed, episodically since 2007, an abstract interpretation framework for security and program securization that is the transformation of a program into a secured program satisfying security criteria defined by a human or artificial supervisor (this is verification when no transformation is needed). The securization is based on the notion of responsibility analysis determining which choices in the program (inputs, random draws, interrupts, schedules, etc.) can definitely cause or avoid desired or menacing events, or have no control at all on the occurrence of these events. Various securization policies (eager, early or late lazy, etc.) have been identified to prevent or enforce the occurrence of events.

6.11. Shape Analysis

We have extended the XISA (eXtensible Inductive Shape Analysis) framework, in order to better deal with low level coding styles and programming languages, and in order to analyze recursive programs in a context dependent way. We also introduced a classification for semantic memory models.

6.11.1. Abstracting Calling-Context with Shapes

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

Interprocedural program analysis is often performed by computing procedure summaries. While possible, computing adequate summaries is difficult, particularly in the presence of recursive procedures. In [23], we propose a complementary framework for interprocedural analysis based on a direct abstraction of the calling context. Specifically, our approach exploits the inductive structure of a calling context by treating it directly as a stack of activation records. We built an abstraction based on separation logic with inductive definitions. A key element of this abstract domain is the use of parameters to refine the meaning of such call stack summaries and thus express relations across activation records and with the heap. In essence, we define an abstract interpretation-based analysis framework for recursive programs that permits a fluid per call site abstraction of the call stack—much like how shape analyzers enable a fluid per program point abstraction of the heap.

6.11.2. Abstract domains for the analysis of programs manipulating complex data-structures

Participant: Xavier Rival.

We proposed a framework for building abstract domains for the static analysis of programs which manipulate complex data-structures [8]. Our abstract domain is parametric in the choice of a numerical abstract domain to represent properties of numeric memory cells and in the choice of a set of inductive definitions to be used in order to summarize unbounded heap regions. It features standard primitives for the computation of transfer functions, for the inclusion checking and for the computation of widening iterates. We also proposed an extension to handle programs that make use of low-level memory addressing, and proposed an extension of the widening to infer inductive definitions.

6.11.3. Composite abstract domain for the analysis of dynamic structures

Participants: Xavier Rival, Antoine Touphans.

Reduced product is a general operation to combine abstract domains into more powerful abstract domains, which has been especially used to construct numerical abstract domains. However, until now, it has not been applied to memory structures. We proposed an instance of a reduced product operation, which can be applied on shape abstract domains based on separation logic and on inductive definitions. The advantage of this construction is that it allows to describe more complex heap dynamic data structures without making the design of all abstract operation more complex. In the other hand, it incurs a reduction cost, whenever we need to transport some information from one domain to the other. We showed that optimal reduction cannot be achieved, and identified the main source of complexity of this operation. A prototype implementation was also carried out. This work was done as part of Antoine Touphans Master internship.

6.12. Static Analysis of Parallel Software

Participant: Antoine Miné.
We present in [22] a static analysis by abstract interpretation to check for run-time errors in parallel C programs. Following our work on ASTRÉE, we focus on embedded critical programs without recursion nor dynamic memory allocation, but extend the analysis to a static set of threads. Our method iterates a slightly modified non-parallel analysis over each thread in turn, until thread interferences stabilize. We prove the soundness of the method with respect to a sequential consistent semantics and a reasonable weakly consistent memory semantics. We then show how to take into account mutual exclusion and thread priorities through partitioning over the scheduler state. We present preliminary experimental results analyzing a real program with our prototype ASTRÉE (see 5.3) and demonstrate the scalability of our approach.

6.13. Termination
Participants: Patrick Cousot, Radhia Cousot.

In [17], we have introduced an abstract interpretation for termination. Proof, verification and analysis methods for termination all rely on two induction principles: (1) a variant function or induction on data ensuring progress towards the end and (2) some form of induction on the program structure.

So far, no clear design principle did exist for termination as is the case for safety so that the existing approaches are scattered and largely not comparable with each other.

For (1), we show that this design principle applies equally well to potential and definite termination. The trace-based termination collecting semantics is given a fixpoint definition. Its abstraction yields a fixpoint definition of the best variant function. By further abstraction of this best variant function, we derive the Floyd/Turing termination proof method as well as new static analysis methods to effectively compute approximations of this best variant function.

For (2), we introduce a generalization of the syntactic notion of structural induction (as found in Hoare logic) into a semantic structural induction based on the new semantic concept of inductive trace cover covering execution traces by segments, a new basis for formulating program properties. Its abstractions allow for generalized recursive proof, verification and static analysis methods by induction on both program structure, control, and data. Examples of particular instances include Floyd’s handling of loop cut-points as well as nested loops, Burstall’s intermittent assertion total correctness proof method, and Podelski-Rybalchenko transition invariants.

6.14. Theories, Solvers and Static Analysis
Participants: Patrick Cousot, Radhia Cousot, Laurent Mauborgne [IMDEA Software (Madrid, Spain)].

In [20], we have introduced a reduced product combining algebraic and logical abstractions to design program correctness verifiers and static analyzers by abstract interpretation. The key new idea is to show that the Nelson-Oppen procedure for combining theories in SMT-solvers computes a reduced product in an observational semantics, so that algebraic and logical abstract interpretations can naturally be combined in a classical way using a reduced product on this observational semantics. The main practical benefit is that reductions can be performed within the logical abstract domains, within the algebraic abstract domains, and also between the logical and the algebraic abstract domains, including the case of abstractions evolving during the analysis.

6.15. Underapproximation for Precondition Inference
Participants: Patrick Cousot, Radhia Cousot, Francesco Logozzo [Microsoft Research (Redmond, USA)], Manuel Fähndrich [Microsoft Research (Redmond, USA)].
In the context of program design by contracts, programmers often insert assertions in their code to be optionally checked at runtime, at least during the debugging phase. These assertions would better be given as a precondition of the method/procedure in which they appear. Potential errors would be discovered earlier and, more importantly, the precondition could be used in the context of separate static program analysis as part of the abstract semantics of the code. However in the case of collections (data structures such as arrays, lists, etc) checking both the precondition and the assertions at runtime appears superfluous and costly. So the precondition is often omitted since it is checked anyway at runtime by the assertions. It follows that the static analysis can be much less precise, a fact that can be difficult to understand since “the precondition and assertions are equivalent” (i.e. at runtime, up to the time at which warnings are produced, but not statically) e.g. for separate static analysis. Moreover preconditions are often understood as overapproximations and thus may exclude good runs which is counter-intuitive for programmers. On the contrary, with considering underapproximations [37], [28] which exclude no good run, ensures that if the precondition is violated then a runtime error must definitely be raised later, and if the precondition is not strong enough to catch all errors they will definitely be captured by a later runtime check.

In [19], we define precisely and formally the contract inference problem from intermittent assertions on scalar variables and elements of collections inserted in the code by the programmer. Our definition excludes no good run even when a non-deterministic choice (e.g. an interactive input) could lead to a bad one. We then introduce new abstract interpretation-based methods to automatically infer both the static contract precondition of a method/procedure and the code to check it at runtime on scalar and collection variables. It has been implemented in CCCHECK (formerly CLOUSOT) by Francesco Logozzo and Manuel Fähndrich.

6.16. Verification of spreadsheet programs by abstract interpretation

Participants: Tie Cheng, Xavier Rival.

Spreadsheet tools (Excel, Openoffice) come with powerful languages which can manipulate sheets in various ways. However, no type discipline is enforced, so that the programs may corrupt spreadsheet contents in many ways. We proposed an abstraction to describe sets of valid spreadsheet states, and designed a verifier for invariants expressed in this abstract domain. Our verifier assumes invariants are defined at the head of loops in the programs (as widening operators for the inference of loop invariants). This work was done as part of Tie Cheng Master internship.
5. New Results

5.1. Analysis of Algorithms

The following articles, conference communications and reports summarize new results in analysis of algorithms over the period: [20], [18], [17], [19], [16], [13], [4], [10], [11], [3], [5].

5.2. Computer Algebra

The following articles, conference communications and reports summarize new results in computer algebra over the period: [9], [21], [2], [8], [22], [7], [15], [1], [12], [6], [14].
6. New Results

6.1. Logical time in Model-Based embedded design

Participants: Charles André, Frédéric Mallet, Julien Deantoni, Robert de Simone, Marie-Agnès Peraldi Frati, Régis Gascon, Calin Glitia, Kelly Garces Pernett, Benoît Ferrero, Nicolas Chleq, Arda Goknil.

The foundational basis of our approach to modeling and analysis of embedded systems using logical time and logical clock specification contraints (CCSL) is recalled in 3.2, and was surveyed in [2]. This year we conducted a number of works exploiting this approach and promoting its introduction to various application domains.

Charles André presented the general approach in an invited lecture at the French Summer School on Real-Time, in Brest [21].

The HDR manuscript of Frédéric Mallet, where the MARTE Time Model is deeply considered, also in relation with other standards such as AADL, was published in book format [39].

In the article [19] we showed how CCSL observers could be encoded in the synchronous language Esterel, using crucial features of simultaneity, and how otherwise simultaneity could be obtained in simulation. This work was also presented internally as deliverable of the FUI Lambda project (see 8.2.3).

We drew a definite link with our activities on Process Network analysis (see 6.3), by showing how the CCSL primitives could be used to provide the loose timing semantic constraints of exiting PN models such as SDF (Synchronous Data-Flow domain of UC Berkeley’s Ptolemy), and its Multi-Dimensiona extension (MD-SDF). This resulted in a journal publication [38]. Existing static schedules can then be obtained by analysis with K-Passa 5.2, or simulated using TimeSquare 5.1 (with an ASAP strategy).

In a collaboration with researchers at East China Normal University (ECNU Shanghai), we showed how CCSL constraints could be translated towards the PROMELA language implemented in the SPIN model-checker, which once again raises the issues of faithfully modeling simultaneity. This work resulted in a communication at the ICECCS conference [33]. Following this work one of our co-author, Yin Ling, earned a one-year scholarship from the Chinese government to visit us as part of her PhD.

The usage of CCSL expressions in the role of predicate property formulas, and their comparison with the more classical temporal formalisms such as PSL (Property Specification Language), was investigated in [24]. A longer internal report version can be found at [42].

In [23] we tackle the issue of recovering global information from multiple execution traces living in distinct logical time bases, with polychronous constraints relating them. The use for efficient debug of embedded systems from distributed traces is examplified on a case study of terrestrial robot. This work was conducted in the framework of the ANR RT-Simex project, see 8.2.1.

A case study in modeling with logical time and CCSL, from requirements to implementation, based on an automotive spark ignition system, is provided in [31]. We worked more generally on the introduction of our approach to existing formalism in the automotive domain, such as EAST-ADL2 and AutoSar, as part of our contribution to the new ITEA2 Timmo2U project. Premises of this effort are described in [32].

The use of CCSL constraints in general requirement engineering was also studied and demonstrated in a conference article, jointly with colleagues at ECNU Shanghai, presented at APSEC’2011 [22].

The use and modeling of priorities amongst timed events (i.e., logical clock ticks), which has strong impacts on efficient logical clock based simulations and scheduling (as the choice of next event), is still a topic of ongoing work. Several advanced considerations are to be found as part of jean-François Le Tallec PhD thesis, to be defended in January 2012 [16].
6.2. Model-Based approaches to SoC design

Participants: Charles André, Robert de Simone, Benoît Ferrero, Carlos Gomez Cardenas, Jean-François Le Tallec.

The main practical goal of this work was to combine in a sensible way the various formalisms SystemC, IP-XACT, UML MARTE, and UPF (for power consumption representation) (see 4.1 for further descriptions). There were true motivations for this: SystemC is a de-facto standard for SoC simulation at various levels, but currently lacks any sort of formal description so that systems can be analysed, reasoned about for correctness and optimized (and it becomes even more so with newer draft standard evolutions). IP-XACT was introduced as an ADL to ease composition and assembly of IP components (written in SystemC or not), but again it currently fails short of its goal, and in particular does not allow standard decoration of model attributes in prominent non-functional domains such as timing/performance and low-power/energy consumption. These could be provided with the help of dedicated features in UML MARTE, and aligned on the UPF standard for power management modeling.

While the intended design flow would take the UML MARTE and UPF to IP-XACT to SystemC direction, it was important to extract IP-XACT and MARTE structural representation from existing SystemC programs, both to populate the flow with existing legacy models, and to explore better the requirements for complete and consistent modeling towards IP block assembly. This work was conducted in Jean-François Le Tallec PhD, to be defended in January 2012 [16]. Together with Benoît Ferrero he defined and realized a software tool named SCiPX (SystemC to IP-XACT translator), originally based on the PinaVM tool by VERIMAG and the DoxyGen syntactic analyzer.

SCiPX is available in prototype version from our site http://www-sop.inria.fr/aoste/index.php?page=software/scipx. It can be combined with the former software transformation modules IPXACT2Marte and Marte2IPXACT developed previously. These results were partly supported by the ID/TL-M contract with ST Microelectronics (see 7.1), and the ANR HeLP project (see 8.2.2), and were presented in [26], [25].

As part of his PhD thesis, Carlos Gomez Cardenas described a subset of UPF standard as a metamodel inside UML MARTE. He also considered compatibility and interconnections with the industrial environments AcePlorer (by Docea Power), and Synopsys Virtualizer (formerly CoWare), provided to us in the context of the CIM PACA tool farm 8.1.1. Preliminary results were presented in [36]. This work was also presented during internal meetings of the ANR HeLP project, and coordinated with work conducted in the team of Michel Auguin at CNRS UMR LEAT (also in Sophia-Antipolis).

6.3. Process Network analysis


This year we comforted the type of analysis on regular static scheduling and routing in dedicated process network models such as studied in the successive PhD thesis of Julien Boucaron, Jean-Vivien Millo, and Anthony Coadou, and recently surveyed in [5]. This resulted mainly in further implementation upgrades of our K-Passa tool (see 5.2), performed first by Anthony Coadou (before he left on a postdoc internship), then continued by Jean-Vivien Millo (on a return postdoc position with us).

In a work mostly conducted while member of the Alchemy EPI in Saclay, but which draws a clear link to our past and present activities in the subject, Sid-Ahmed-Ali Touati studied efficient heuristics to the general problem of one-dimensional periodic task scheduling under storage requirements, using a modeling framework akin to Process Networks. This resulted in a journal article accepted once the author had become attached to the Aoste EPI [17].

6.4. Correct and efficient implementation of polychronous formalisms

Participants: Thomas Carle, Manel Djemal, Virginia Papailiopoulou, Dumitru Potop Butucaru, Robert de Simone, Yves Sorel.
Existing analysis techniques for synchronous and polychronous languages, such as clock calculi, are meant to extract relations of simultaneity (time inclusion) and exclusiveness (time exclusion) between the various computations and communications. This approach is well-suited when targeting sequential processors. For distributed or multi-threaded implementations, further independence relations are needed to express potential concurrency. This resulted in a general theory of endochronous systems, meant to support this additional analysis [11].

Last year we completed a first prototype tool implementation for weak endochrony checking. This was completed this year in two directions:

- connecting our tool with Signal as input language, and interface it in practice to the Polychrony/SME environment developed by the Espresso EPI;
- Improving algorithmic complexity and internal data representaion, so that our tool can now handle reasonable size Signal programs.

This work was of course conducted in collaboration with Espresso members. Experimental results were presented at the ESLsyn 2011 conference [30]. We are currently expanding the framework in order to take modes/states into account in the program specifications. Effective generation of multi-threaded GALS wrappers for Signal programs is also under way.

We worked at extending the AAA methodology for polychronous processes by providing a better integration of clock analysis in the various phases of the implementation process (allocation, scheduling, pipelining, etc.). We also considered a wider range of implementation targets (time-triggered, MPSoC). We defined a dedicated software pipelining algorithm to match conditional scheduling/reservation tables such as used in SynDEx, with the goal of improving throughput with the same duration of individual computation cycles (as is the goal of any pipelining techniques). The originality here is to make logical clocks of polychronous systems act as triggers for the predicated executions as used in traditional software pipelining. First results have been presented during the Synchron 2011 workshop and in a research report [41].

Further work on time-triggered systems was submitted inside the FUI Parsec 8.2.4 and P 8.2.5 projects, including real-time implementation methods for the IMA/ARINC 653 avionics platforms. In particular we conducted experiments to replace the scheduling policy of the second-level scheduler (L1 in the standard) from dynamic priority-driven to dynamic Time Division time triggered (TT-IMA). Preliminary results are under way, and were informally presented at the yearly Synchron seminar.

An important emerging trend in target MPSoC platforms is that On-chip networks are progressively introduced to cope with the bottleneck of inter processor communications. Correct implementation of polychronous systems in this context thus relies on efficient routing of data in such networks, and ultimately may assume that on-chip NoC routers may be programmed in one way or another to behave predictably according to the global application distributed on the cores. We started a collaboration on this topic with the "Embedded Systems-on-Chips" department of the LIP6 laboratory, one of the main site of expertise for SoC/NoC design and Hardware/software codesign. This collaboration first materialized with the co-supervision of M. Djemal’s PhD thesis. A generic MPSoC architecture is being defined, which includes a 2D mesh network-on-chip with programmable routers, on which static routing schedules such as synthesized by our tools may be implemented and run.

### 6.5. Uniprocessor Real-time Scheduling

**Participants:** Laurent George, Mohamed Marouf, Daniel de Rauglaudre, Yves Sorel.

#### 6.5.1. Strict periodic harmonic tasks

This year, we focused our work on scheduling of strict periodic tasks to the particular case of harmonic tasks [28]. After transforming the scheduling problem into a bin-packing problem, we performed a schedulability analysis and proposed schedulability conditions in each sub-case of harmonic tasks: we proposed a necessary and sufficient condition in the case where all tasks periods are distinct, and we proposed a sufficient condition in the case where some tasks have the same period. Finally, we proposed a scheduling algorithm based on the bin-packing problem resolution.
6.5.2. Combination of strict periodic and sporadic tasks

Non-preemptive strict periodic tasks are harder to schedule than preemptive ones. One can hope to extend schedulability results when combining non-preemptive strict periodic tasks with preemptive sporadic one.

We proposed in [27] a schedulability analysis for a combination of strict periodic and sporadic tasks. We considered all tasks with fixed priorities, where the highest priorities are given to strict periodic tasks and the lower priorities are given to sporadic tasks. First, we scheduled strict periodic tasks using our former scheduling algorithm. Then, we computed the critical instants which maximize the response time of a sporadic task. We proved that the critical instants are contained in the permanent phase of strict periodic tasks, and are given by the start times of strict periodic jobs in a hyper-period. We also proved that we can reduce critical instants by eliminating some of them. Then, we gave the analytic expression of the computing time $W_i(t)$ at any time $t$ necessary for the execution of a task $\tau_i$ taking in consideration all the tasks with higher priorities. That allows the computation of the response time $r_i$ by solving the equation $W_i(t) = r_i$. Therefore, for a sporadic task, if its response time $r_i$ is less of equal to its deadline for all critical instants, then this latter task is schedulable, else it is not schedulable. We proceed similarly for all sporadic tasks to prove that a tasks set is schedulable or not.

6.5.3. Exact cost of RTOS

It is important to determine the exact cost of the real-time operating system (RTOS) when preemptive scheduling is used for better processor utilization compared to non preemptive scheduling [43]. Indeed, in this case it is possible to trust the schedulability conditions when they include that cost, and also to avoid waste resources. This year we developed a generic RTOS modelled with Petri nets and we determined its exact cost on an ARM9 processor. We used Petri nets on the one hand to choose through simulations the best structure of that scheduler, and on the other hand to verify non blocking properties. In order to obtain its exact cost the scheduler was programmed in assembly language, and coded such as it is deterministic, i.e. its cost does not depend on alternative statements but only on the number of tasks which is known a priori. Using this RTOS we experimented simple task sets on the ARM9 processor for which we were able to include the exact RTOS cost in the schedulability conditions.

6.5.4. Formal proofs of real-time scheduling theorems

Scheduling involves numerous models and theorems, sometimes dated of several decades, but never formally proved. We made a formal proof in Coq (proof assistant developped at Inria) to check a classical theorem giving a schedulability condition for a set of real-time strictly periodic tasks (about 1500 lines of Coq). This work was published in a paper accepted for publication in the conference JFLA 2012.

A second proof is actually being carried now, dealing with response time of a set of fixed priority real-time preemptive tasks. The theorem states that the worst case of this response time occurs when all tasks start simultaneously. A step in the original argument by Jane W. S. Liu [55] involves the proof of a function whose fixpoint computes the response time of the first instance of the least priority task. This specific step is now formally proved in Coq (3500 lines of Coq), and we are now working on the completion of the full theorem.

6.6. Multiprocessor Real-time Scheduling

Participants: Laurent George, Maxence Guesdon, Mohamed Marouf, Falou Ndoye, Simon Nivault, Yves Sorel, Cécile Stentzel.

6.6.1. Partitioned scheduling with exact RTOS cost

In the case of partitioned scheduling we propose a greedy heuristic to solve the real-time scheduling problem of periodic preemptive tasks on a multiprocessor architecture while taking into account the exact RTOS cost. This is achieved by combining an allocation heuristic, of “best fit” type, and a schedulability condition based on the operation $\oplus$ which takes into account the exact RTOS cost [43]. The allocation heuristic minimizes the makespan (total execution time of the tasks executed on the multiprocessor taking into account inter-processor communication costs). A first version of that work was presented in [29].
6.6.2. Semi-partitioned scheduling

In [18] we study two cases of semi-partitioned scheduling of sporadic tasks with constrained deadlines on homogeneous multi-processor: (i) the case where the Worst Case Execution Time (WCET) of a job can be portioned, each portion being executed on a dedicated processor, according to a static pattern of migration and using for solving the local assignment problem a linear programming approach; (ii) the case where the jobs of a task are released on a processor, 1 time out of $p$, where $p$ is an integer less than or equal to the number of processors, according to a Round-Robin migration pattern. The first approach has been investigated in the state-of-the-art by migrating a job at its local deadline, computed from the deadline of the task it belongs to.

6.6.3. Fault tolerance on electric vehicles

We consider applications composed of a real-time task set running on the distributed heterogeneous architecture of the CyCab (electric vehicle developed in the IMARA team-project) based on dsPICs processors, MPC555 micro-controllers, and an embedded PC all together connected through CAN (Controller Area Network) buses. For hardware reasons we suppose that only dsPICs and CAN buses can fail. Our goal is to find a fault-tolerant software solution to tolerate such failures while the applications satisfy the real-time constraints. Because extra hardware for error detection is expensive in such electric vehicle, we proposed a software error detection based on watchdogs. We solved separately two different problems: buses and dsPICs fault-tolerance. In both cases we use active redundancy policies. For buses fault-tolerance, we assume that all processors are reliable, and all but one bus can fail. The same data is sent through all the CAN buses. If a CAN bus fails then the data is sent by the other CAN buses. For processors fault-tolerance, we assume that all communication media are reliable and at least one processor can fail. The first step consists in performing active redundancy for all the tasks of the application. A task and their redundant tasks are assigned to different processors. If processor fails then the data which are not sent by tasks running on that faulty processor, are actually sent by the redundant tasks. All the tasks with their redundant counterparts are scheduled according to the schedulability analysis proposed in [28].

6.6.4. Scicos/SynDEx gateway and code generation for multi-core

This work was carried out in the Openprod project (see 8.3.2.2). The gateway between Scicos and SynDEx has been updated to deal with the last Scicos data structures and the last version of SynDEx. Besides, this gateway has been improved and partially rewritten to support as much Scicos blocks as possible. We use the gateway to automatically produce from a control model specified and simulated in Scicos a real-time executable running on a multi-core platform. The latter platform is described according to the shared memory model defined last year. In order to generate real-time executable code we had to develop a new SynDEx executive kernel based on Windows-RTX which supports shared memory communications and multi-core parallel execution. That executive kernel is used with the macro-code generated by SynDEx to produce the real-time executable code.

6.6.5. SynDEx updates

We continued the developments of future version 8 of SynDEx which will feature a new software architecture to allow better functionality evolutions and maintenance. On the other hand in the COTROS ADT ("Génération de code temps réel distribué optimisé et sûr"), we achieved the new automatic code generator for the current version 7 of SynDEx. This generator creates intra and inter-processor synchronizations according to well defined rules, checks deadlock absence in inter-processor synchronizations, manages efficiently buffers and semaphores (declaration, naming, etc.).
CASCADE Project-Team (section vide)
6. New Results

6.1. Constraint Handling Rules and Linear Logic

Participants: François Fages, Thierry Martinez, Sylvain Soliman.

Implementations of Constraint Handling Rules (CHR) follow a committed-choice forward chaining execution model: the non-determinism of the abstract semantics is partly refined with extra-logical syntactic convention on the program order and possibly notations for weighted semantics (with priorities or probabilities), and partly left unspecified in the underlying compiler. In [13], we propose an alternative execution model which explores all the possible choices, by opposition to the committed-choice strategy. This execution model is angelic in the sense that if there exists a successful execution strategy (with respect to a given observable), then this strategy will be found. Formally, the set of computed goals is complete with respect to the set of the logical consequences of the interpretation of the initial goal in linear logic. In practice, this paper introduces a new data representation for sets of goals, the derivation nets. Sharing strategies between computation paths can be defined for derivation nets to make execution algorithmically tractable in some cases where a naive exploration would be exponential. Control for refined execution is recovered with the introduction of user constraints to encode sequencing, fully captured in the linear-logic interpretation. As a consequence of angelic execution, CHR rules become decomposable while preserving accessibility properties. This decomposability makes natural the definition in angelic CHR of meta-interpreters to change the execution strategy. More generally, arbitrary computation can be interleaved during head matching, for custom user constraint indexation and deep guard definition.

6.2. Rule-based Modeling Language for Constraint Programming

Participants: François Fages, Raphaël Martin, Thierry Martinez, Sylvain Soliman.

Rules2CP is a rule-based modeling language which allows easy modeling of constraint satisfaction problems, together with specifications for search strategies and heuristic choice criterias by pattern matching. In [23], we study a new compilation scheme for Rules2CP which allows us to deal with dynamic ordering criteria and to generate procedural constraint programming code instead of flattened constraints. The comparison with the static expansion of Rules2CP models shows that the overhead at runtime is limited, with a gain in the size of the generated program which could be exponentially larger by static expansion.

The language Rules2CP is currently extended to deal with hybrid discrete and continuous domains and packing problems with complex shapes, in the framework of the ANR Net-WMS-2 project with KLS-Optim and EMN Inria EPI TASC. The compiler of Rules2CP is currently rewritten in Java in the framework of a collaboration with KLS-Optim supported by Inria DTI.

6.3. Trace Development Methodology


We are working on a general theory of traces design taking traces as primary objects of study. It is based on the observation of the way trace files are accumulated as knowledge bases and elaborated in different fields of activity like software engineering, rule based systems and resolution, learning in context, or personal experience storing systems.

We worked on two main points: the development of an experimental tracer of CHR with TODAS project and an application the notion of generic trace to standardization of constraints. In [17] we analyze, and occasionally correct, shortcomings of the former approach based on the generic trace format GENTra4CP, and show the interest that a generic tracer may bring to develop portable applications or to standardization efforts, in particular in the field of constraints.
6.4. Railway Time Tabling Optimization

**Participants:** François Fages, David Fournier, Sylvain Soliman.

Sustainable development is a key issue for our society. Optimization of resources, energy and costs (as admirably done in living organisms) has thus grown significantly over years to become a major field in industry. In collaboration with General Electric Transportation France which is a key-actor in the field of transportation all around the world, we investigate energy reduction for train and metro service providers through time tabling optimization. In [31], we describe and compare different optimization methods to reduce energy in mass rapid transits (MRT). Most of the literature deals with a special problematic arising in train services: the maximum traction energy. They show that for reducing costs and energy consumption, one method is the reduction of the peak energy over a time period. This objective function has been chosen for this study and the thesis depicts how to implement it on different paradigms, such as mixed-integer linear programming, constraint programming or local search. We conclude on promising approaches in terms of optimization methods for time tabling computation and real-time scheduling.

6.5. Petri Net Analysis of Biochemical Networks

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

Bridging the gap between quantitative and qualitative models, Petri nets (also known as place/transition graphs) have recently emerged as a promising tool for modeling and analysis of biochemical networks. In [14], we present a method to compute the minimal siphons and traps of a Petri net as a Constraint Satisfaction Problem (CSP). In our case, siphons and traps are purely structural properties that brings us information about the persistence of some molecular species. We present a constraint program that finds minimal siphons and traps containing specific set of places in a Petri net. This method is compared on models of the biomodels.net repository with other methods based on Mixed Integer Linear Programming (MILP) and Boolean Satisfiability (SAT). The flexibility brought by constraint programming, for instance in the declarative choice of variable enumeration heuristics, seems promising in further improving those results.

6.6. Theory of Subgraph Epimorphisms

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

The operations of deleting and merging vertices are natural operations for reducing a graph. While graph reductions through a sequence of vertex deletions (resp. mergings) characterize subgraph isomorphisms (resp. graph epimorphisms), sequences of both vertex deletion and merging operations characterize subgraph epimorphisms. Our proposal is thus to use subgraph epimorphism for comparing graphs in applications where a more flexible notion than the classical notion of subgraph isomorphism is required.

In collaboration with Christine Solnon (INSA Lyon), we have developed the theory of subgraph epimorphisms in [5]. We have shown that SEPIs preserve graph completeness and arc symmetry and that, just like SISO and EPI, SEPI is not a well quasi order. We have defined the SEPI, EPI and SISO distances between two graphs as the size of the largest SEPI (resp. EPI, SISO) lower bound graphs. These distances are equal to the minimum number of respectively vertex deletion and/or merging operations that are necessary to obtain isomorphic graphs. They are also metrics on graphs and we have $d_d \geq d_{md}$ and $d_m \geq d_{md}$.

From a computational point of view, we have shown that the existence of a SEPI between two graphs is an NP-complete problem and have presented a constraint satisfaction algorithm for solving it. This algorithm is implemented in BIOCHAM.

It is worth noticing that, given two graphs $G$ and $G'$, the greatest lower SEPI bounds and the least upper SEPI bounds are also interesting to compute since they represent “intersection” and “union” graphs for the SEPI relation. For instance, in our motivating application in systems biology, these objects correspond to the intersection (resp. union) of models at different levels of details for a given biochemical process. These graphs are not unique but we are confident that the constraint satisfaction algorithm presented in [10] can be interestingly generalized to compute them.
6.7. Parameter Search under Temporal Logic Constraints

Participants: Grégory Batt, Elisabetta De Maria, François Fages, Domitille Heitzler, Aurélien Rizk, Sylvain Soliman, Jannis Uhlendorf.

Our method for solving temporal logic constraints in the quantifier-free fragment of first-order linear time logic QFLTL(R), opens up the field of model-checking to optimization through the definition of a continuous degree of satisfaction for temporal logic formulae [8], [2]. This satisfaction degree can be used in a number of ways, e.g. as a fitness function with continuous optimization methods to find unknown parameter values in a model with temporal logic constraints formalizing biological properties [4], [6], or to control a system from a temporal specification of its behavior [15], or to compute the robustness of a system w.r.t. a temporal property and a perturbation of the parameters.

This approach is implemented in BIOCHAM and is one unique feature of this modeling environment.

6.8. Model-based Optimization of Cancer Chronotherapies

Participants: Elisabetta De Maria, François Fages, Aurélien Rizk, Sylvain Soliman, Denis Thieffry.

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

In [4] we show how temporal logic constraints, and the new features of BIOCHAM for parameter search (running on a cluster of 10000 processors at the GENCI) can be used to couple dynamical models in high dimension and more precisely to build a coupled model composed of:

- a four phases model of the mammalian cell cycle by Novak and Tyson,
- a circadian clock model by Leloup and Goldbeter,
- a DNA damage repair model by Ciliberto et al.,
- a model of irinotecan metabolism by Dimitrio and Ballesta,
- a simple model of drug administration control.

This coupled model allows us to minimize the toxicity of irinotecan on healthy cells, using BIOCHAM’s parameter search method applied on the drug administration control law.

Our technology is ready to calibrate models on real patient data, evaluate model predictions and optimize patient-tailored chronotherapeutics. The collaboration currently focuses on the obtaining of consistent data in the C5Sys project and on the improvement of the cell cycle model.

6.9. Analysis of FSH and Angiotensine Signaling

Participants: François Fages, Domitille Heitzler, Aurélien Rizk, Sylvain Soliman.

In [6] in collaboration with Eric Reiter (UMR CNRS-INRA 6175) and Frédérique Clément (SISYPHE) in the framework of the Initiative Action REGATE, we have combined experimental approaches with computational modeling to decipher the molecular mechanisms as well as the hidden dynamics governing ERK activation by the angiotensin II type 1A receptor (AT1AR) in HEK293 cells. We have built in BIOCHAM a dynamical model that captures available knowledge and experimental data. The unknown kinetic parameters have been inferred using a temporal logic specification of experimental data in both control and perturbed conditions, using a cluster of 10000 processors at the GENCI.

\(^2\)we use the Covariance Matrix Adaptation Evolutionary Strategy CMAES of Nikolaus Hansen from the EPI TAO. Moreover, this year we have implemented a second method by Particle Swarm Optimization, PSO.
The mathematical model predicts and experiments confirm that, for the AT1AR expressed in HEK293 cells:
i) GRK2/3 and 5/6 regulate switching between the G protein and β-arrestin pathways as well as their distinct
dynamics by phosphorylating the C-terminal region of the activated receptor; ii) GRK2/3 not only mediates
desensitization of G protein activation but also exerts a strong restraining influence on β-arrestin signaling;
iii) GRK5/6 exert little effect on G protein-stimulated ERK but are required for β-arrestin-mediated ERK
activation; iv) the β-arrestin-dependent ERK pathway undergoes both activation and deactivation through
amplified enzymatic processes.

These results convincingly illustrate the value of using computational modeling to decipher the complex
signaling mechanisms elicited by 7TMRs [1]. This approach is applied more generally to G protein-coupled
receptor signaling which is of great importance in pharmacology.

6.10. Multi-affine Hybrid Automaton Model of Cardiac Cells

Participant: Grégory Batt.

A fundamental question in the treatment of cardiac disorders, such as tachycardia and fibrillation, is under
what circumstances does such a disorder arise? To answer to this question, in collaboration with E. Bartocci
and R. Grosu at SUNY, Stony Brook, USA, we develop a multi-affine hybrid automaton (MHA) cardiac-cell
model, and restate the original question as one of identification of the parameter ranges under which the MHA
model accurately reproduces the disorder [12]. The MHA model is obtained from the minimal cardiac
model of one of the authors, Fenton from Cornell University, by first bringing it into the form of a canonical
genetic regulatory network, and then linearizing its sigmoidal switches in an optimal way. By leveraging the
Rovergene tool for genetic regulatory networks, we are then able to successfully identify the parameter ranges
of interest.

6.11. Real-time Control of Gene Expression in Yeast

Participants: Grégory Batt, François Fages, Jannis Uhlendorf.

To decipher the dynamical functioning of cellular processes, the method of choice is to observe the time
response of cells subjected to well controlled perturbations in time and amplitude. Efficient methods, based
on molecular biology, are available to monitor quantitatively and dynamically many cellular processes. In
contrast, it is still a challenge to perturb cellular processes - such as gene expression - in a precise and controlled
manner. In collaboration with Pascal Hersen at MSC lab (Paris Diderot University), in the framework of the
Iceberg ANR project, we propose a first step towards in vivo control of gene expression: in real-time, we
dynamically control the activity of a yeast signaling cascade thanks to an experimental platform combining
a micro-fluidic device, an epi-fluorescence microscope and software implementing control approaches [15].
We experimentally demonstrate the feasibility of this approach, and we investigate computationally some
possible improvements of our control strategy using a model of the yeast osmo-adaptation response fitted to
our data.

6.12. Artificial Tissue Homeostasis in Mammalian Cells

Participants: Grégory Batt, François Bertaux, Xavier Duportet, François Fages, Szymon Stoma.

Cell-based gene therapy aims at creating and transplanting genetically-modified cells into a patient in order
to treat a disease. Ideally, actively-growing cells are used to form a self-maintaining tissue in the patient, thus
permanently curing the disease. However, before any real therapeutic use, robust mechanisms enforcing tissue
homeostasis, that is, that the size of the newly-introduced tissue remains within admissible bounds, need to be
developed. We proposed various designs and tested their robustness using in silico approaches. Preliminary
results demonstrated that cell-to-cell variability plays a crucial role for tissue long-term maintenance. More
extensive in silico characterizations require the development of efficient multiscale simulation methods. In
parallel to the in silico work, done in collaboration with the Bang research group (Dirk Drasdo), we started the
construction and in vitro experimental characterization of the most promising designs in collaboration with
the Weiss lab (MIT) [24].
FORMES Team

6. New Results

6.1. Simulation

6.1.1. Simulation of vector architecture

Participants: Vania Joloboff, Yang Yu.

Many architectures including PowerPC and ARM now have vectorized instructions, that is, instructions that can execute on several data items in parallel (e.g. 8 simultaneous additions) on specific vector data.

We have implemented the ALTIVEC extension of the PowerPC to support the vector instructions.

6.1.2. Native translation using LLVM

Participants: Vania Joloboff, Xinlei Zhou, Zuyu Zhang.

We have started to implement a new technique of dynamic translation. This new method consists in decompiling the binary object code into an abstract representation and recompiling it to native host code.

The decompilation of the program amounts to reconstructing the simulated program Control Flow Graph using an intermediate representation. We have chosen LLVM (Low Level Virtual Machine), defined by University of Illinois, and now widely adopted in many projects, as our representation format. Using LLVM allows us to directly use the LLVM Intel code generator.

The SimSoC binary decoder has been modified to identify basic blocks (blocks of sequential instructions ending with a branch instruction). After instructions have been grouped into basic blocks, they are translated into an LLVM representation and finally the LLVM compiler is called to generate native code.

A first version of this technique has been implemented for both the ARM and Power Architecture. We have reach a considerable speed improvement in the generated code, with the execution speed multiplied by factor of 2 to 8. However the translation time from binary to LLVM and from LLVM to native code is significant (translation speed is roughly 1000 instructions per second). Consequently the overall speed is improved by only a factor of 20 to 50 percent when the simulation are relatively short test programs [20].

In order to reach still higher simulation speed we need to use a more sophisticated analysis of the control flow graph. The idea is to do an edge profiling analysis of the basic blocks in order to identify larger blocks. This work is under development.

Another idea is to use multi-processor hosts machine to parallelize translation from LLVM to native code. This is also under investigation.

6.1.3. Trace Analysis

Participants: Guillaume Merle, Vania Joloboff.

Simulation sessions produce huge trace files, sometimes now in hundreds of gigabytes, that are hard to analyze with a quick response time. This comes down to two sub-problems:

- The trace file size. Trace files are huge because they include lots of information. But when looking for a specific problem, one does not need all of this information. To search one given defect, one may ignore a large amount of the data in the trace file. One would like the trace file to contain only relevant information to the concerned problem.

- The expressive power of the language to analyze the trace, and its usability. If the language is limited to expression search, it is easy to use but hard to construct sophisticated formulas. If the language used is Linear Temporal Logic (LTL), there is a lot of expressive power but many engineers are unable to write a LTL formula and to maintain it over time.
We would like to build a trace analysis tool that includes a language which allows expression of time-related formulas but is simple to formulate expressions. When this language is compiled, ideally the compiler is smart enough to identify independent formulae, the search of which can be parallelized, and it is also smart enough to generate "filter scripts".

When compiling one trace language input file, it would generate, from one input file, N filter scripts and N analyzers. Then during the simulation, the huge raw trace file is actually split into N smaller trace files, each relevant to one problem only, filtering out all unnecessary data. Hence trace files sizes would be considerably reduced.

We have started to design a trace language and a compiler, and extended the SimSoC simulator to support generation of trace files with a filter.

A first version of the trace language compiler has been coded in OCAML.

In the current version under development, the filters are not generated but coded manually, and filters are not parallelized.

6.1.4. Generation of simulators from vendor specification

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

Starting last year, we undertook the task of generating automatically an instruction set simulator (ISS) from the vendor specification in a PDF file. In order to generate the C code of the simulator, it is assumed such vendor specification contains at least some formal definitions of the instruction set that can be analyzed. It is the case to a wide extent for the ARM, the PowerPC and the SH architectures.

The process of generating the simulator consists of 4 major steps, first eliminating from the PDF file irrelevant information, next construct from the relevant data an abstract syntax representation of the instruction set, then to generate the C code of the simulator, using some additional data provided manually to complete the vendor specification.

This work was completed last year for the ARM architecture with the documentation form ARM corporation [35]. This year, we did similar work for the SH architecture from specification from RENESAS corporation. We have indeed generated a simulator for the SH4 architecture [31], which has not been fully tested yet.

However, this works has proved that the abstract syntax we have defined is powerful enough to describe two different architectures with significant differences in the way they are described by the vendor.

6.1.5. First steps towards the certification of an ARM simulator

Participants: Frédéric Blanqui, Jean-François Monin, Xiaomu Shi, Frédéric Tuong.

The simulation of Systems-on-Chip (SoC) is nowadays a hot topic because, beyond providing many debugging facilities, it allows the development of dedicated software before the hardware is available. Low-consumption CPUs such as ARM play a central role in SoC. However, the effectiveness of simulation depends on the faithfulness of the simulator. To this effect, in [24], we propose here to prove significant parts of such a simulator, SimSoC. Basically, on one hand, we develop a Coq formal model of the ARM architecture while on the other hand, we consider a version of the simulator including components written in Compcert-C [61]. Then we prove that the simulation of ARM operations, according to Compcert-C formal semantics, conforms to the expected formal model of ARM. Size issues are partly dealt with using automatic generation of significant parts of the Coq model and of SimSoC from the official textual definition of ARM. However, this is still a long-term project. We report here the current stage of our efforts and discuss in particular the use of Compcert-C in this framework.

6.2. Type and rewriting theory

6.2.1. A type theory for Coq

Participants: Jean-Pierre Jouannaud, Qian Wang.
In this joint work with Bruno Barras and Pierre-Yves Strub [17], we describe an abstract model of CoqMT [73] called CoqMTU, which puts together the Calculus of Inductive Constructions, decidable first-order theories, and an infinite hierarchy of universes which are all predicative but the first impredicative universe of propositions. We have shown its consistency, strong normalization and decidability of type checking in presence of weak elimination (and absence of strong elimination). An important feature of this work is that the first-order theory is abstract, characterized by the three natural axioms that (i) it is non-degenerated (its models have at least two elements), (ii) constructors are free, and (iii) defined symbols are completely defined. On the theoretical side, this allows us to give an abstract elimination principle for such non-canonical theories. On the practical side, this justifies the implementation of CoqMT in which decidable theories can be dynamically downloaded. It should be noticed that these proofs are done in Coq, except for the strong normalization part. Qian Wang is now continuing this work at Ecole Polytechnique with Bruno Barras and Pierre-Yves Strub, the target being strong normalization.

6.2.2. Confluence by decreasing diagrams
Participants: Jean-Pierre Jouannaud, Huiying Luo, Jiaxiang Liu.

Invented by Vincent Van Oostrom, decreasing diagrams capture both kinds of diagrams arising from Newman’s Lemma and Hindley’s Lemma: they indeed allow to reduce all known confluence methods to critical pairs computations, and a search of decreasing diagrams for them all, where decreasingness is measured by a well-founded order on proof steps.

In [55], we give a new simple proof of Van Oostrom’s main theorem, and extend the method of decreasing diagrams to rewrite relations on a term algebra. We prove that the union of a terminating left-linear systems, and a non-terminating linear system is confluent provided the various critical pairs existing in in their combination have decreasing diagrams (with respect to some order built from the respective orders of both systems).

During this year, we have further simplified and generalized these results in order to get rid of the left-linearity assumption for the first system, and of the right-linearity assumption for the second. This yields a true generalization of the well-known Knuth-Bendix-Huet confluence result for terminating systems, and at the same time of various critical-pair based results found in the literature for non-terminating systems.

6.2.3. Confluence of normal rewriting
Participants: Jean-Pierre Jouannaud, Jianqi Li.

Confluence results for first-order and higher-order rewriting differ in many ways: by the rewriting relation used, and by the strong normalization assumption made. We believe that these differences hide the strong similarities of these (and other) kinds of rewriting.

In this work, we introduce a new notion of rewriting, normal rewriting, which aims at capturing all known results reducing confluence to critical (and extension) pair computations in presence of some termination assumption.

We achieve this goal in the following way. First, we consider theories made of a set $R$ of rules, a set $S$ of simplifiers, and a set $E$ of equations. Rewriting operates on terms in $S$ modulo $E$ normal forms, and uses $S \cup E$-pattern matching for firing the rules in $R$, before to normalize the result with respect to $S$ modulo $E$. Termination is assumed for the union of $S$ modulo $E$ and $R$ modulo $S \cup E$. Second, we introduce relations on an abstract set of terms, and an abstract, well-founded set of positions, and reduce the Church-Rosser property of abstract normal rewriting to abstract notions of critical pairs and extensions. We can then apply this result to first-order rewriting, as well as to various forms of higher-order rewriting. These results capture plain rewriting ($S \cup E = \emptyset$), Stickel’s rewriting modulo ($S = \emptyset$), Nipkow’s higher-order rewriting ($S$ is made of beta-reduction and eta-expansion, and $E$ is alpha-conversion), and allow to describe new forms of first and higher-order rewriting relations.

6.2.4. Argument filterings and usable rules in higher-order rewrite systems
Participant: Frédéric Blanqui.
Joint work with Keiichirou Kusakari and Sho Suzuki from Nagoya University, Japan.

The static dependency pair method is a method for proving the termination of higher-order rewrite systems à la Nipkow [62]. It combines the dependency pair method introduced for first-order rewrite systems with the notion of strong computability introduced for typed lambda-calculi [52]. Argument filterings and usable rules are two important methods of the dependency pair framework used by current state-of-the-art first-order automated termination provers [51], [53]. In [12], we extend the class of higher-order systems on which the static dependency pair method can be applied. Then, we extend argument filterings and usable rules to higher-order rewriting, hence providing the basis for a powerful automated termination prover for higher-order rewrite systems.

6.3. Decision procedures

6.3.1. A certificate framework for DPLL(T)

Participants: Min Zhou, Fei He, Bow-Yaw Wang, Wenrui Meng.

Satisfiability Modulo Theories (SMT) techniques are widely used nowadays. SMT solvers are used to decide the satisfiability of first-order formulas. When an SMT solver is invoked, it is important to ensure correctness of the result. For this purpose, we proposed a certificate framework based on DPLL(T), including generation of certificates and verification of certificates. Some properties are discussed and proved theoretically. The certificate is easy to generate because it only needs minor modification to the existing SMT solvers. Experiment results show that the overhead for certificates generation is only 10%. Moreover, verifying the certificate requires few memory and time, which outperforms other approaches.

6.3.2. Automated verification of termination certificates

Participants: Frédéric Blanqui, Kim-Quyen Ly, Sidi Ould Biha.

The research community on rewriting developed a grammar for termination certificates called CPF [29] (given by a XML Schema file). Our goal is to develop a safe, modular and efficient termination certificate verifier based on the formal library of mathematical results on termination called CoLoR that has been developed for the proof assistant Coq [11].

Because the CPF format is regularly modified and extended with new features, it is useful to have a tool that can automatically generate data structures, parsers and pretty-printers for that format. Hence, we developed a first version of such a tool in OCaml.

Once we got a representation of termination certificates in Coq, we could start defining a boolean function checking the correctness of a certificate, and formally prove its correctness. For the moment, we only considered the case of polynomial interpretations on integers. The proof is almost finished. To do so, we had to modify some of the CoLoR files to be able to use its results (transformation of modules into records that are first-class objects). The use of dependent types in CoLoR makes also definitions and proofs much more difficult.

6.3.3. Proving computational geometry algorithms in TLA+2

Participants: Hui Kong, Hehua Zhang, Ming Gu.

Geometric algorithms are widely used in many scientific fields like computer vision, computer graphics. To guarantee the correctness of these algorithms, it is important to apply formal method to them. In this work, we propose an approach to proving the correctness of geometric algorithms [22]. The main contribution is that a set of proof decomposition rules is proposed which can help improve the automation of the proof of geometric algorithms. We choose TLA+2, a structural specification and proof language, as our experiment environment. The case study on a classical convex hull algorithm shows the usability of the method.

6.4. Compositional verification

6.4.1. BDD-based assume-guarantee reasoning through implicit learning

Participants: Fei He, Bow-Yaw Wang, Lei Zhu.
We present a purely BDD-based assume-guarantee reasoning technique to improve the scalability of symbolic model checking. The new technique adopts a BDD learning algorithm to generate BDD’s as contextual assumptions. A new witness analysis algorithm is proposed to exploit the multitude of traces returned by symbolic model checkers. Using the classification tree-based BDD learning algorithm to generate contextual assumptions, we compare assume-guarantee reasoning with monolithic symbolic model checking. The new technique always infers smaller contextual assumptions than contexts in our experiments.

6.4.2. Predicate generation for learning-based loop invariant inference


We address the predicate generation problem in the context of loop invariant inference. Motivated by the interpolation-based abstraction refinement technique, we apply the interpolation theorem to synthesize predicates implicitly implied by program texts. Our technique is able to improve the effectiveness and efficiency of the learning-based loop invariant inference algorithm in [21]. Experiments excerpted from Linux, SPEC2000, and Tar source codes are reported.

This is a joint work with Yungbum Jung, Wonchan Lee, and Kwangkuen Yi of Seoul National University, South Korea.

6.4.3. Thread-modular model checking with iterative refinement

Participants: Wenrui Meng, Fei He, Bow-Yaw Wang.

Thread-modular analysis is an incomplete compositional technique for verifying concurrent systems. The heuristic works rather well when there is limited interaction among system components. In this project, we develop a refinement algorithm that makes thread-modular model checking complete. Our algorithm refines abstract reachable states by exposing local information through auxiliary variables. The experiments show that our complete thread-modular model checking can outperform other complete compositional reasoning techniques.

6.5. Specification and verification of TLA+ and PLC systems

6.5.1. Formal semantics of PLC programming languages

Participants: Sidi Ould Biha, Litian Xiao, Ming Gu.

We formalized a semantics of the Instruction List (IL) language, one of the five programming languages defined in the IEC 61131-3 standard for PLC programming [23]. This semantics support a significant subset of the IL language that includes on-delay timers. This semantics was used in a join work to with Jan Olaf Blech from Fortiss (Germany) to prove some safety properties for a real industrial example of PLC program [18].

A second widely used language for programming PLC is the graphical language Ladder Diagrams (LD). We defined a formal semantics of LD in the proof assistant Coq. Based on this semantics and the IL one, we developed a translation function from LD to IL. We also proved a semantic preservation property for this translation function. We have now a certified compilation function from the graphical language LD to IL. This work opens the way for the development of a certified compilation chain for PLC. A journal paper about this work and others aspects of PLC certification is under reviewing.

In [16], [15], we study the definition of denotational semantics on PLC program language, which is convenient to PLC programs modeling and model checking. The purpose of the work is the correctness verification on PLC programs by formal methods. Based on the extended λ-calculus definition, this work has defined the configuration of PLC program architecture, denotational semantics of PLC programs and functions of denotational semantics. It is the basis of model checking and theorem proving.

6.5.2. Formalization and verification of PLCs

Participants: Hai Wan, Litian Xiao, Ming Gu.
PLCs are widely used in embedded systems. Timers play a pivotal role in PLC real-time applications. The formalization of timers is of great importance. In [13], we present a formalization of PLC timers in the theorem proving system Coq, in which the behaviors of timers are characterized by a set of axioms at an abstract level. The authors discuss how to model timers at a proper and sound abstract level. PLC programs with timers are modeled. As a case study, a quiz machine problem with a timer is investigated. This work demonstrates the complexity of formal timer modeling.

In [25], we modeled kernel data type and basic statements and the denotational semantics of PLC program in Coq. It has given the correctness proof of PLC program based on theorem proving, i.e. based on semantics function the relationship of configuration between the before codes execution and the after is proved. The main purpose is to prove whether a PLC program satisfies certain nature within a scan period.

6.5.3. Synthesis of PLC programs

Participants: Rui Wang, Ming Gu.

PLCs are complex cyber-physical systems which are widely used in industry. In [14], we present a robust approach to design and implement PLC-based embedded systems. Timed automata are used to model the controller and its environment. We validate the design model with resort to model checking techniques. We propose an algorithm to generate PLC code from timed automata and implement this algorithm with a prototype tool. This method can condense the developing process and guarantee the correctness of PLC programs. A case study demonstrates the effectiveness of the method.

6.5.4. Domain-driven probabilistic analysis of PLCs

Participants: Hehua Zhang, Yu Jiang, Ming Gu.

Programmable Logic Controllers are widely used in industry. Reliable PLCs are vital to many critical applications. We present a novel symbolic approach for analysis of PLC systems [27]. The main components of the approach consists of: (1) calculating the uncertainty characterization of the PLC systems, (2) abstracting the PLC system as a Hidden Markov Model, (3) solving the Hidden Markov Model using domain knowledge, (4) integrating the solved Hidden Markov Model and the uncertainty characterization to form an integrated (regular) Markov Model, and (5) harnessing probabilistic model checking to analyze properties on the resultant Markov Model. The framework provides expected performance measures of the PLC systems by automated analytical means without expensive simulations. Case studies on an industrial automated system are performed to demonstrate the effectiveness of our approach.

6.5.5. Edola: a domain modeling and verification language for PLCs

Participants: Hehua Zhang, Ming Gu.

Formal modeling and verification of PLC systems become paramount in engineering applications. The work presents a novel PLC domain-specific modeling language Edola [26]. Important characteristics of PLC embedded systems, such as reactivity, scan cycling, real-time and property patterns, are embodied in the language design. Formal verification methods, such as model checking and automatic theorem proving, are supported in Edola modeling. The TLA+ specification language constitutes an intermediate language layer between Edola and the verification tools, enhancing a large degree of reusability. A prototype IDE for Edola and its seamless integration of a model checker TLC and an automatic theorem prover Spass are implemented. A case study illustrates and validates the applicability of the language.

6.6. Distributed algorithms

6.6.1. Formal model and proofs for Netlog protocols

Participants: Meixian Chen, Jean-François Monin.

Joint work with Yuxin Deng (Jiaotong University, Shanghai) and Stéphane Grumbach (LIAMA/Netquest).
Netlog is a language designed and implemented in the Netquest project for describing protocols. Netlog has a precise semantics, provides a high level of abstraction thanks to its Datalog flavor and benefits from an efficient implementation. This makes it a very interesting target language for proofs of protocols. Netlog comes with two possible semantics: a synchronous semantics, better suited to tightly coupled parallel systems and an asynchronous semantics, better suited to distributed systems.

We designed a formal model of Netlog in Coq, where the two possible semantics are derived from common basic blocks. In a fully certified framework, a formal proof of the Netlog engine (running on each node) would be required. We don’t attack this part at the moment: we assume that the implementation respects the general properties stated in our model and focus on the issues raised by the distributed model of computation provided by Netlog.

As a proof of concept, we applied in 2010 this framework to an algorithm constructing a Breadth-First Search Spanning Tree (BFS) in a distributed system [46]. This work has been slightly improved this year and published in [19].

Moreover, we generalized the model in order to take the removal of datalog facts into account, and started to use this feature for more complicated protocols. In main one under study is Prim’s algorithm (publication under submission), and we target next GHS, which still resists to palatable proof techniques.
6. New Results

6.1. Formal verification of compilers and static analyses

6.1.1. The Compcert verified compiler for the C language

Participants: Xavier Leroy, Sandrine Blazy [project-team Celtique], Alexandre Pilkiewicz.

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 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 ways:

- The formal semantics for the CompCert C source language was made executable and turned into a reference interpreter. This interpreter is proved sound and complete with respect to the formal semantics. It makes it possible to animate the semantics on test programs, identifying undefined behaviors and enumerating all possible execution orders. Another application is to provide an experimental validation of the semantics itself.
- The top-level statements of compiler correctness were strengthened. In particular, semantic preservation is shown to hold even in the presence of a non-deterministic execution context. Also, we showed that if the source program goes wrong after performing some input/output actions, the compiled code performs at least these actions before continuing with an arbitrary behavior.
- A new optimization pass, redundant reload optimization, was added, improving performance by up to 10% on the x86 architecture.
- A general annotation mechanism was added to observe the values of local program variables at user-specified program points, such observations being guaranteed to produce the same results in the source code and in the compiled code. These annotations can be used to improve the precision of worst-case execution time (WCET) analysis over the compiled code. They can also provide stronger evidence of traceability for code qualification purposes.

Three versions of the Compcert development were publically released, integrating these improvements: versions 1.8.1 in March, 1.8.2 in April, and 1.9 in August.

In parallel, we continued our collaboration with Jean Souyris, Ricardo Bedin França and Denis Favre-Felix at Airbus. They are conducting an experimental evaluation of CompCert’s usability for avionics software, and studying the regulatory issues (DO-178 certification) surrounding the potential use of CompCert in this context. Preliminary results were reported at the Predictability and Performance in Embedded Systems workshop [20]. More detailed results will be presented at the 2012 Embedded Real-Time Software and Systems conference (ERTS’12) [19].

6.1.2. Formal specification and verified compilation of C++

Participants: Tahina Ramananandro, Gabriel Dos Reis [Texas A&M University], Xavier Leroy.
This year, under Xavier Leroy’s supervision and with precious C++ advice from Gabriel Dos Reis, Tahina Ramananandro tackled the issue of formally specifying object construction and destruction in multiple-inheritance languages, especially the C++ flavour featuring non-virtual and virtual inheritance (allowing repeated and shared base class subobjects), and also structure array fields. This formalization consists in specifying, in Coq, a small-step operational semantics for a subset of C++ featuring multiple inheritance, static and dynamic casts, field accesses, and object construction and destruction, and mechanically proving properties about resource management, thus obtaining a formal account of the RAII (Resource Acquisition is Initialization) principle. Moreover, this formalization also studies the impact of object construction and destruction on the behaviour of dynamic operations such as virtual function dispatch, introducing the notion of generalized dynamic type. These results were accepted for publication at the POPL 2012 symposium [29].

Finally, this formalization includes a verified realistic compiler for this subset of C++ to a CFG-style 3-address intermediate language featuring low-level memory accesses in the style of the CompCert RTL language. Following usual compilation schemes and techniques inspired from the Common Vendor ABI for Itanium (which has since been reused and adapted by GNU GCC), the target language additionally features virtual tables to model object-oriented features, and virtual table tables to model the generalized dynamic type changes during object construction and destruction. This verified compiler reuses and extends the results of a previous work on verified C++ object layout by Tahina Ramananandro, Gabriel Dos Reis and Xavier Leroy published this year at the POPL 2011 symposium [28].

6.1.3. Validation of polyhedral optimizations

Participants: Alexandre Pilkiewicz, François Pottier.

The polyhedral representation of loop nests with affine bounds is a unified way to compute and represent a large set of optimizations, including loop fusion, skewing, splitting, peeling, tiling etc. Polyhedral optimizers usually rely on heavily optimized C tools and libraries to manipulate polyhedrons. Those C libraries are, like any other programs, bug prone, which can easily lead to erroneous optimizations.

Those two facts—powerful yet error prone—make the formal proof of such optimizations appealing. Proving a full optimizer however would probably be unrealistic: the proof would be terribly challenging, but even writing in Coq an optimizer efficient enough to handle non trivial loop nest might be impossible.

Another option is to write and prove in Coq a validator: after each run of the unproved optimizer—considered as a black box—the validator is used to compare the program before and after optimization to make sure that its semantics—the meaning of the program—has not been change. If the validator does not report an error, we have formal certitude that no bug has been introduced by the optimization.

Alexandre Pilkiewicz, under François Pottier’s supervision, has implemented and proved in Coq such a validator.

6.1.4. A formally-verified parser for CompCert


During a 6-month Master’s internship (M2), Jacques-Henri Jourdan built a formally-verified parser for the C99 language. This parser was obtained through a general method for checking that an LR(1) parser produced by the parser generator Menhir is correct and complete, that is, it conforms exactly to the specification represented by the context-free grammar. This check is carried out by a validator that is implemented in Coq and proved correct, so that, in the end, there is no need to trust Menhir. A paper describing this work was accepted for presentation at the ESOP 2012 conference [24].

6.1.5. Formal verification of an alias analysis

Participants: Valentin Robert, Xavier Leroy.
As part of his 5-month Master’s internship, Valentin Robert developed and proved correct a static analysis for pointers and non-aliasing. This alias analysis is inaprocedural and flow-sensitive, and follows the “points-to” approach of Andersen [40]. An originality of this alias analysis is that it is conducted over the RTL intermediate language of the CompCert compiler: since RTL is essentially untyped, the traditional approaches to field sensitivity do not apply, and are replaced by a simple but effective tracking of the numerical offsets of pointers with respect to their base memory blocks. Using the Coq proof assistant and techniques inspired from abstract interpretation, Valentin Robert proved the soundness of his alias analysis against the operational semantics of RTL.

6.2. Type systems

6.2.1. A type-and-capability calculus with hidden state

Participants: François Pottier, Jan Schwinghammer [Saarland University, Saarbrücken], Lars Birkedal [IT University of Copenhagen], Bernhard Reus [University of Sussex, Brighton], Kristian Støvring [University of Copenhagen], Hongseok Yang [University of Oxford].

During the year 2010, François Pottier developed a machine-checked proof of an expressive type-and-capability system. Such a system can be used to type-check and prove properties of imperative ML programs. The proof, which follows a “syntactic” method, is carried out in Coq and takes up roughly 20,000 lines of code. It confirms that earlier publications by Charguéraud and Pottier [1], [7] were indeed correct, offers insights into the design of the type-and-capability system, and provides a firm foundation for further research. In the first half of 2011, François Pottier wrote a paper that describes the system and its proof in detail. This paper has been submitted for publication [37].

Together with Jan Schwinghammer and other co-authors, François Pottier also worked on a (pencil-and-paper) proof of this type-and-capability system. This proof is based on a “semantic” method and is quite different from the proof mentioned in the previous paragraph. It offers somewhat different insights, and proves (for the first time) that the ideas presented in an unpublished note by Pottier (“Generalizing the higher-order frame and anti-frame rules”, 2009) were correct. A paper that describes this proof has been submitted for publication [39].

6.2.2. Fine-grained static control of side effects in HaMLet

Participants: Jonathan Protzenko, François Pottier.

In the past ten years, the type systems community and the separation logic community, among others, have developed highly expressive formalisms for describing ownership policies and controlling side effects in imperative programming languages. In spite of this extensive knowledge, it remains very difficult to come up with a programming language design that is simple, effective (it actually controls side effects!) and expressive (it does not force programmers to alter the design of their data structures and algorithms). Jonathan Protzenko and François Pottier have recently made significant progress on this topic. They are designing a programming language, tentatively called HaMLet, in the tradition of ML and Caml-Light. The language offers immutable and mutable algebraic data structures and first-class functions. It allows very fine-grained control of ownership and side effects. The project is still at a preliminary stage and no publications have appeared yet.

6.2.3. Partial type inference with first-class polymorphism

Participants: Didier Rémy, Boris Yakobowski [CEA, LIST laboratory], Gabriel Scherer.

The language MLF uses optional type annotations of function parameters and instance bounded polymorphism—quantification over all types that are instances of a given type—to smoothly combine the simple type inference mechanism of ML with the expressive types of System F. In MLF, programs need only type annotations on parameters of functions that are used polymorphically in their body.

While the surface language requires just these very few type annotations, MLF also comes with an internal language, called xMLF, where all type manipulations become explicit so that they can traced during program transformations and symbolic evaluation. The internal language is described in a journal paper [13].
Gabriel Scherer has maintained and improved a prototype implementation of MLF including the elaboration of MLF into xMLF and an extension to higher-order types.

### 6.2.4. First-class module systems

**Participants:** Benoît Montagu [University of Pennsylvania], Didier Rémy, Gabriel Scherer.

Singleton kinds are used to handle type definitions in modules. They accurately model the propagation of type definitions through higher-order functor applications. However, type equivalence in the presence of singleton types is hard to formalize and to implement. In his PhD dissertation [48], Benoît Montagu has proposed a new way of checking equivalence in the presence of singleton types, based on expensors. Expensors are eta-expansion constants that are inserted in the source program in such a way that equivalence of two programs becomes equality of their normal forms after insertion of expensors. This approach was described in an article to be submitted to a conference.

Since October, Gabriel Scherer has been working on mixin modules. Mixin modules are an attractive generalization of modules with horizontal composition, a mechanism that allows more flexible construction of modules.

Gabriel Scherer has been studying whether the use of open existential types introduced earlier by Benoît Montagu for first-class modules can be used to simplify the presentation of mixin modules, hoping that they could be given a direct semantics, instead of one by means of elaboration into another language with recursive modules.

### 6.2.5. Coercion abstraction

**Participants:** Julien Cretin, Didier Rémy.

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

We studied F_π, a language where all type transformations are represented as coercions. This language provides polymorphism as in System F, (upper) bounded polymorphism as in F_{<:}, lower bounded polymorphism as in MLF, and η-expansion as in F_η. Hence, F_π unifies these four languages in a generic framework.

We showed that F_π has a type erasing semantics by bisimulation with the lambda calculus. This means that coercions can be dropped before evaluation without changing the meaning of programs.

This work is described in a paper to be presented at the POPL 2012 conference [21] and in a technical report [33].

### 6.2.6. Kind-level typing in Haskell

**Participants:** Julien Cretin, Brent Yorgey [University of Pennsylvania], Stephanie Weirich [University of Pennsylvania], José Pedro Magalhães [Utrecht University], Simon Peyton Jones [Microsoft Research Cambridge], Dimitrios Vytiniotis [Microsoft Research Cambridge].

Haskell is a functional programming language with a rich static type system. Programmers use advanced type features to enforce invariants over data structures. This quickly leads to the need for computation in types. Until now, computation at the type level was untyped in Haskell and therefore prone to errors and hard to debug.

We extended the kind level of Haskell with two features already present at the type level: data types and polymorphism. These features are already well-known at the type level, and should remain easy to understand for programmers at the kind level.

Kind polymorphism is now implemented and used in the core language of the
Glasgow Haskell Compiler (GHC). Promotion of data-types is implemented in a branch of GHC. Both extensions are described in a paper to be presented at the TLDI 2012 workshop [31].

6.3. Software specification and verification

6.3.1. Proved time complexity bounds for program components

Participants: Sylvain Dailler, François Pottier.

During a six-month master internship (M2), Sylvain Dailler extended Arthur Charguéraud’s CFML tool with a notion of “time credit”. This allows CFML to be used to prove not only that an algorithm (or a data structure, or a library) is correct, but also that it meets a desired worst-case asymptotic complexity bound. Because CFML is hosted within Coq, these proofs are machine-checked. Sylvain Dailler was able to establish the functional correctness and the time complexity of a library that implements “bags” as circular doubly-linked lists [35].

6.3.2. Hybrid contract checking via symbolic simplification

Participant: Na Xu.

Program errors are hard to detect or prove absent. Allowing programmers to write formal and precise specifications, especially in the form of contracts, is one popular approach to program verification and error discovery. Na Xu formalized and implemented a hybrid contract checker for a subset of OCaml. The key technique is the use of symbolic simplification, which makes integrating static and dynamic contract checking easy and effective. This technique statically verifies that a function satisfies its contract or blames the function violating the contract. When a contract satisfaction is undecidable, it leaves residual code for dynamic contract checking. A paper describing this result will be presented at the PEPM’2012 conference [30]. A technical report version is also available [34].

6.3.3. Tools for TLA+

Participants: Damien Doligez, Leslie Lamport [Microsoft Research], Stephan Merz [EPI VeriDis], Denis Cousineau [Microsoft Research-INRIA Joint Centre], Markus Kuppe [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 [45], and to build tools for writing TLA+ specifications and mechanically checking the corresponding formal proofs.

This year, the TLA+ project prepared the release of the third version of the TLA+ tools: the GUI-based TLA Toolbox and the TLA+ Proof System, an environment for writing and checking TLA+ proofs. This new release will add many improvements in terms of efficiency, notably with a system of fingerprints to support incremental development of proofs. It will also bring support for new back-ends based on SMT provers (CVC3, Z3, Yices, VeriT). This extends the range of proof obligations that the system can discharge automatically.

Web site: http://tlaplus.net/.

6.3.4. The Zenon automatic theorem prover

Participant: Damien Doligez.

Damien Doligez continued the development of Zenon, a tableau-based prover for first-order logic with equality and theory-specific extensions. This year, a refactoring of the prover’s architecture was started.

6.4. The Caml language and system

6.4.1. The OCaml system

Participants: Xavier Clerc [team SED], Damien Doligez, Alain Frisch [Lexifi SAS], Jacques Garrigue [University of Nagoya], Fabrice Le Fessant [EPI Asap and OCamlPro start-up company], Jacques Le Normand [Lexifi SAS], Xavier Leroy, Nicolas Pouillard, Pierre Weis [EPI Estime].
This year, we released version 3.12.1 of the OCaml system. This is a minor release that fixes 65 reported bugs and 9 unreported bugs, and introduces 11 small extensions. Damien Doligez acted as release manager for this version.

In parallel, we have been working on the next major release of OCaml. The major innovation is support for generalized algebraic datatypes (GADTs). These non-uniform datatype definitions enable programmers to express some invariants over data structures, and the OCaml type-checker to enforce these invariants. They also support interesting ways of reflecting types into run-time values. GADTs are found in proof assistants such as Coq and in functional languages such as Agda and Haskell. Their integration in OCaml raised delicate issues of partial type inference and principality of inferred types, to which Jacques Garrigue and Jacques Le Normand provided original solutions [43].

Other features in preparation for the next major release include:

- More lightweight first-class modules. Signature annotations over first-class modules can now be omitted when they are determined by the context.
- Better reporting of type errors: shorter but more relevant context is shown; improved tracking of source code locations in modules.
- Improvements in native-code generation, for instance in the case of partial function applications.
- Improvements in the generic hashing primitive and the standard library for hash tables.

6.4.2. Customizable unmarshaling for OCaml

Participants: Pascal Cuq [CEA LIST], Damien Doligez, Julien Signoles [CEA LIST].

In collaboration with members of the CEA LIST laboratory, Damien Doligez developed a Caml library for treating marshaled data by applying user-specified on-the-fly transformations during the unmarshaling process. This library is used in CEA's Frama-C software to support marshaling of hash-consed data. The library was presented at the ML workshop [32].

6.5. Meta-programming

Participants: Nicolas Pouillard, François Pottier.

In an effort to improve meta-programming support (the ability to write programs that manipulate other programs) in programming languages, we have focused first on the issue of binders. Programming with data structures containing binders occurs frequently: from compilers and static analysis tools to theorem provers and code generators, it is necessary to manipulate abstract syntax trees, type expressions, logical formulae, proof terms, etc. All these data structures contain variables and binding constructs.

Nicolas Pouillard, under the supervision of François Pottier, investigated the design of a programming interface for names and binders where the representations of these two types are kept abstract. This interface is sufficiently general to enable a large body of program transformations.

This year, the de Bruijn indices approach has been investigated more in-depth, resulting in a programming interface specialized to safe programming with de Bruijn indices and providing much more precise results than those published in 2010. This work was published at the ICFP 2011 conference [27].

6.6. Formal management of package dependencies

Participants: Roberto Di Cosmo, Ralf Treinen [University Paris Diderot], Jaap Boender [University Paris Diderot], Pietro Abate [University Paris Diderot], Jérôme Vouillon [University Paris Diderot], Stefano Zacchiroli [University Paris Diderot].

Roberto Di Cosmo’s current main line of research is the study and analysis of large component-based software repositories, in particular GNU/Linux-based distributions. These distributions consists of collections of dozens of thousands of software packages, together with metadata, installation and configuration tools, and a variety of different production processes, involving quality assurance at several levels.
Ensuring quality of software assemblies built using these components is a challenging issue: the simple question of knowing whether a single component can or not be deployed turns out to be NP-complete, and yet industry needs to deploy components all the time.

The research currently conducted within the Mancoosi FP7 european project, coordinated by Roberto Di Cosmo, addresses some of the relevant issues, by elaborating sophisticated deployment algorithms and designing specialised installation and configuration languages targeted at enabling transactional capabilities in the tools used to maintain software assemblies built out of GNU/Linux based distributions.

The results of this project are available at http://www.mancoosi.org/ and include four publications this year: one at the CBSE conference [18], which received an ACM distinguished paper award; one at the FSE conference [23], which received an ACM distinguished artifact award; one in the Science of Computer Programming journal [12]; and one at the workshop on Logics for Component Configuration [22].
6. New Results

6.1. Weak Memory Models, Litmus-PPC-WMM tools

Participants: Jade Alglave, Luc Maranget, Pankaj Pawan, Susmit Sarkar [U. of Cambridge], Peter Sewell [U. of Cambridge], Francesco Zappa Nardelli.

Shared memory multiprocessors typically expose subtle, poorly understood and poorly specified relaxed-memory semantics to programmers. To understand them, and to develop formal models to use in program verification, we find it essential to take an empirical approach, testing what results parallel programs can actually produce when executed on the hardware. We describe a key ingredient of our approach, our litmus tool, which takes small “litmus test” programs and runs them for many iterations to find interesting behaviour. It embodies various techniques for making such interesting behaviour appear more frequently. We presented a tool, litmus, to run “litmus tests” on real hardware at TACAS’11 [31].

An automated exploration of machine memory models (on the dont tool) is submitted to TACAS’12.

During a two month long internship, Pankaj Pawan (IIT Kanpur, India) ported the PPCMEM application from OCaml to JavaScript, and developed a suitable web-interface. This enabled a wide dissemination (including at IBM) of the PPCMEM tool. The tool is available online at: http://www.cl.cam.ac.uk/~pes20/ppcmem/help.html.

6.2. Operational semantics for the memory model of Power (IBM) machines

Participants: Jade Alglave, Luc Maranget, Susmit Sarkar [U. of Cambridge], Peter Sewell [U. of Cambridge], Derek Williams [IBM, Austin].

Exploiting today’s multiprocessors requires high-performance and correct concurrent systems code (optimising compilers, language runtimes, OS kernels, etc.), which in turn requires a good understanding of the observable processor behaviour that can be relied on. Unfortunately this critical hardware/software interface is not at all clear for several current multiprocessors.

We characterise the behaviour of IBM POWER multiprocessors, which have a subtle and highly relaxed memory model (ARM multiprocessors have a very similar architecture in this respect). We have conducted extensive experiments on several generations of processors: POWER G5, 5, 6, and 7. Based on these, on published details of the microarchitectures, and on discussions with IBM staff, we give an abstract-machine semantics that abstracts from most of the implementation detail but explains the behaviour of a range of subtle examples. Our semantics is explained in prose but defined in rigorous machine-processed mathematics; we also confirm that it captures the observable processor behaviour, or the architectural intent, for our examples with an executable checker. While not officially sanctioned by the vendor, we believe that this model gives a reasonable basis for reasoning about current POWER multiprocessors. Our work should bring new clarity to concurrent systems programming for these architectures, and is a necessary precondition for any analysis or verification. It should also inform the design of languages such as C and C++, where the language memory model is constrained by what can be efficiently compiled to such multiprocessors.

This work was presented at PLDI’11 [26]. This operational model is now being enriched for Power machine with loads and link/store conditionals (the Power primitives to write locks), and connected to C++ semantics. It is submitted to PLDI’2012 with many co-authors from Cambridge.

6.3. Restoring Sequential Consistency for x86 and Power machines

Participants: Jade Alglave, Luc Maranget.
Concurrent programs running on weak memory models exhibit relaxed behaviours, making them hard to understand and to debug. To use standard verification techniques on such programs, we can force them to behave as if running on a Sequentially Consistent (SC) model. Thus, we examine how to constrain the behaviour of such programs via synchronisation to ensure what we call their stability, i.e. that they behave as if they were running on a stronger model than the actual one, e.g. SC. First, we define sufficient conditions ensuring stability to a program, and show that Power’s locks and read-modify-write primitives meet them. Second, we minimise the amount of required synchronisation by characterising which parts of a given execution should be synchronised. Third, we characterise the programs stable from a weak architecture to SC. Finally, we present our offence tool which places either lock-based or lock-free synchronisation in an x86 or Power program to ensure its stability.

This work was presented at CAV’11 [30]

6.4. Relaxed-Memory Concurrency and Verified Compilation


We studied the semantic design and verified compilation of a C-like programming language for concurrent shared-memory computation above x86 multiprocessors. The design of such a language is made surprisingly subtle by several factors: the relaxed-memory behaviour of the hardware, the effects of compiler optimisation on concurrent code, the need to support high-performance concurrent algorithms, and the desire for a reasonably simple programming model. In turn, this complexity makes verified (or verifying) compilation both essential and challenging.

This project started in 2010, when we defined a concurrent relaxed-memory semantics for ClightTSO, an extension of CompCert’s Clight in which the processor memory model is exposed for high-performance code, and, building on CompCert, we implemented a certifying compiler from ClightTSO to x86, and proved in Coq several compilation phases. A paper describing our approach has been accepted in POPL [29] 2011.

During 2011 we completed this project by developing correctness proofs for all the compilation phases, and made a public distribution of the compiler, available from http://www.cl.cam.ac.uk/~pes20/CompCertTSO.

In 2011 Zappa Nardelli and Vafeiadis investigated the soundness of fence elimination optimisations for x86TSO. They implemented and proved correct two optimisations that remove redundant fence instructions as compiler passes over RTL in CompCertTSO. Despite an apparent simplicity, these optimisations generate almost optimal code for several standard concurrent algorithms (CompCertTSO does not implement escape analysis, which would enhance the effectiveness of the optimisations), and since they only perform data-flow analysis over the code of each thread (without analysing the full-system thread interactions) they do not suffer from the finite-state and finite control limitation of other approaches. The proof of correctness of one optimisation was challenging has required some for prophecy variable simulation. This work has been published in SAS 2011 [28] and the code is part of CompCertTSO.

A journal version, describing the correctness proof of all the phases of CompCertTSO (including the fence eliminations) as been submitted to the Journal of the ACM [35].

6.5. Compiling C/C++ Concurrency: from C++11 to POWER

Participants: Kayvan Memarian, Francesco Zappa Nardelli.

The upcoming C and C++ revised standards add concurrency to the languages, for the first time, in the form of a subtle relaxed memory model (the C++11 model). This aims to permit compiler optimisation and to accommodate the differing relaxed-memory behaviours of mainstream multiprocessors, combining simple semantics for most code with high-performance low-level atomics for concurrency libraries.

We studied the the correctness of two proposed compilation schemes for the C++11 load and store concurrency primitives to Power assembly, having noted that an earlier proposal was flawed. (The main ideas apply also to ARM, which has a similar relaxed memory architecture.)
This should inform the ongoing development of production compilers for C++11 and C1x, clarifies what properties of the machine architecture are required, and builds confidence in the C++11 and Power semantics. A paper describing this work will appear in POPL 2012 [22].

6.6. F*: Secure Distributed Programming with Value-Dependent Types

Participants: Nikhil Swamy [MSR Redmond], Juan Chen [MSR Redmond], Cédric Fournet [MSR Cambridge], Pierre-Yves Strub [MSR-INRIA], Karthikeyan Bhargavan [correspondant], Jean Yang [MIT].

Distributed applications are difficult to program reliably and securely. Dependently typed functional languages promise to prevent broad classes of errors and vulnerabilities, and to enable program verification to proceed side-by-side with development. However, as recursion, effects, and rich libraries are added, using types to reason about programs, specifications, and proofs becomes challenging.

We present F*, a full-fledged design and implementation of a new dependently typed language for secure distributed programming. Unlike prior languages, F* provides arbitrary recursion while maintaining a logically consistent core; it enables modular reasoning about state and other effects using affine types; and it supports proofs of refinement properties using a mixture of cryptographic evidence and logical proof terms. The key mechanism is a new kind system that tracks several sub-languages within F* and controls their interaction. F* subsumes two previous languages, F7 and Fine. We prove type soundness (with proofs mechanized in Coq) and logical consistency for F*.

We have implemented a compiler that translates F* to .NET bytecode, based on a prototype for Fine. F* provides access to libraries for concurrency, networking, cryptography, and interoperability with C#, F#, and the other .NET languages. The compiler produces verifiable binaries with 60% size overhead for proofs and types, as much as a 45x improvement over the Fine compiler, while still enabling efficient bytecode verification.

To date, we have programmed and verified more than 20,000 lines of F* including (1) new schemes for multi-party sessions; (2) a zero-knowledge privacy-preserving payment protocol; (3) a provenance-aware curated database; (4) a suite of 17 web-browser extensions verified for authorization properties; and (5) a cloud-hosted multi-tier web application with a verified reference monitor.

This paper was published at ICFP 2011 [27].

6.7. Security by Typing for Cryptographic Protocol Implementations

Participants: Karthikeyan Bhargavan [correspondant], Cédric Fournet [MSR Cambridge], Andrew D. Gordon [MSR Cambridge], Alfredo Pironti.

We propose to use refinement typing to verify the security of cryptographic protocol implementations. Our method is based on declaring and enforcing invariants on the usage of cryptography. We develop cryptographic libraries that embed a logic model of their cryptographic structures and that specify preconditions and postconditions on their functions so as to maintain their invariants.

We implement the method for protocols coded in F# and verified using F7, our SMT-based typechecker for refinement types, that is, types carrying formulas to record invariants. As illustrated by a series of programming examples, our method can flexibly deal with a range of different cryptographic constructions and protocols [24].

We are currently evaluating this method on a fully-fledged implementation of TLS. While previous uses of typing for cryptographic protocol implementations focused on the symbolic model of cryptography, we use a new technique by Fournet et al to develop computational proofs for our implementations. Our TLS implementation consists of 6000 lines of code. We have currently annotated and verified about half of this implementation.

We recently published a tutorial on our verification method as part of the proceedings of FOSAD 2010, and a journal paper on our type system at TOPLAS [20].
6.8. Authorization for the Social Web: from Formal Analysis to Concrete Attacks

Participants: Chetan Bansal [BITS Goa], Karthikeyan Bhargavan [correspondent], Sergio Maffeis [Imperial College].

Social sign-on and social sharing are becoming an ever more popular feature of web applications. This success is caused in part by the APIs support offered by leading websites such as Facebook, Twitter and Google, and by the openness of standards such as OAuth 2.0. A formal analysis of such protocols must account for malicious websites and their JavaScript, and common website vulnerabilities, such as cross site request forgery and open redirectors.

We present a formal model for web application protocols called WebSpi, implemented as a library for the protocol verification tool ProVerif. We use WebSpi to model and verify several configurations of the OAuth 2.0 protocol. We show that our formal analysis can be used to reconstruct concrete website attacks. Our approach is validated by finding dozens of previously unknown vulnerabilities in popular websites such as Yahoo and Wordpress, when they connect to social networks such as Twitter and Facebook.

We are in discussion with Facebook, Twitter and other websites to address the vulnerabilities found by our analysis. We have submitted a paper describing this work, and plan to release the WebSpi library in 2012.

6.9. Verified Android Cryptographic Applications

Participants: Karthikeyan Bhargavan [correspondent], Quentin Lefebvre [MPRI].

With the emergence of application markets for smartphones, hundreds of third-party applications now use cryptography to protect sensitive user data before storing it on disk or sending it out on the network. However, using cryptographic mechanisms correctly to fulfill a desired security goal is challenging and error-prone, even for experts. Our goal is to build verification tools that developers may use to develop security proofs for their applications.

We show how to verify the security of third-party cryptographic applications written in Java for the Android platform. We first develop symbolic security specifications for classes in the JCA. We can then verify that applications that use these libraries satisfy their security goals, even in the presence of a Dolev-Yao adversary who controls the network, the disk, and potentially other applications on the device. We report preliminary verification results using the Krakatoa verification tool for Java programs.

We presented this work at the ASA workshop 2011 [23].

6.10. Logical Auditing of JavaScript Programs for Security

Participants: Karthikeyan Bhargavan [correspondent], Sergio Maffeis [Imperial College], Ravinder Shankesi [UIUC].

Client side web applications are error-prone and hard to secure, as proven by frequent vulnerability reports. We experiment with using logical annotations as a means to specify inlined security policies for web pages, and we implement a run-time monitoring system that generates a logical trace of the program execution. Feeding the logical trace to external theorem provers, it is possible to detect, post-facto, violations of the security policies, helping the on-line debugging of web applications.

We present JSTY a browser-based logical auditing framework for JavaScript programs. We show how first-order logic contracts can be used to express cryptographic assumptions and security goals for JavaScript programs that use cryptography. We demonstrate our approach on realistic examples, including browser extensions for password management. We find security vulnerabilities in commercial products by logical auditing. This work is currently unpublished.

6.11. Secure Interpreters for Sessions

Participants: James Leifer, Jean Pichon [intern from ENST Telecom ParisTech].
We developed an interpreter for decentralised multi-party sessions. The interpreter takes a graph-based description of a session and provides a high-level interface for sending and receiving messages permissible in the session. The interpreter protects the integrity of session execution in a realistic security setting where an adversary has the ability to: (1) control the network to capture and reinject messages at will, and read and forge messages using leaked cryptographic keys; (2) compromise arbitrary session participants. By producing and verifying cryptographic signatures, the interpreter ensures that all non-compromised participants have consistent views of the session’s execution history.

We previously worked on a session compiler. The compiler took as input “local graphs” described in process-calculus fashion. It produced one F# library for each role, and an F7 interface/specification for this library. If the library typechecked (with the F7 type-checker) against its interface, then it was secure. However, the production of the library was not itself verified, and could fail.

By contrast, this present work is concerned with a session interpreter. This interpreter works like an ML functor (i.e. a compile-time function from modules to modules). It takes as input a module describing a global graph by exposing a specific interface. The application of the interpreter functor to a graph module yields an F$\#$ module that contains the interface/specification. The interpreter functor, being checked against an abstract description of a graph, is typechecked “once and for all”. Because the interpreter functor typechecks against its interface (that contains specifications), it is secure (through refinements). The interpreter functor can then be applied to a concrete session graph to be used.

The interpreter is written in F$\#$, a dialect of ML enriched with refinement types, and its correctness is proven by type annotations.

The interpreter consists of approximately 3000 lines of code.

### 6.12. Models of Audit Logs

**Participants:** Karthikeyan Bhargavan, Cédric Fournet [MSR Cambridge], Nataliya Guts, Francesco Zappa Nardelli.

This line of research was accurately described in last year activity report of Moscova. Here we just mention that Nataliya Guts defended her PhD [19] on "Auditability for security protocols" on January 11th, 2011.
6. New Results

6.1. Compilation techniques for synchronous languages

- The paper *Modular Static Scheduling of Synchronous Data-flow Programs* by M. Pouzet and P. Raymond (VERIMAG Grenoble) has been selected among the two best papers at EMSOFT 2009. An extended version is published in a special issue of the *Journal of Design Automation for Embedded Systems*, in 2010. This work solves a 20 years problem raised by P. Raymond in 88 [59].

- Significant extensions have been provided to the Heptagon language. C. Pasteur contributed a memory optimization pass, combining a static analysis with user-given annotations (a paper on this subject is submitted to LCTES’2012), and the integration of discrete controller synthesis, developed and used by the SARDES team at INRIA Rhones-Alpes. Other work-in-progress extensions and experiments conducted inside the PARKAS team include the generation of VHDL code, parallel code generation, and n-synchronous code generation from Lucy-n.

6.2. Semantics and Implementation of Hybrid System Modelers

Hybrid systems modelers have become the corner stone of embedded system development, with Simulink 11 a de facto standard and Modelica 12 a new player. They allow both discrete controllers and their continuous environments to be expressed in a single language. Despite the availability of such tools, there remain a number of issues related to the lack of reproducibility of simulations and to the separation of the continuous part, which has to be exercised by a numerical solver, from the discrete part, which must be guaranteed not to evolve during a step. Such tools still raise a number of issues that, we believe, require more fundamental understanding.

In collaboration with Albert Benveniste and Benoit Caillaud (INRIA Rennes) we have proposed using non standard analysis as a semantic domain for hybrid systems. Non standard analysis is an extension of classical analysis in which infinitesimals can be manipulated as first class citizens. This allows us to provide a denotational semantics and a constructive semantics for hybrid systems, thus establishing simulation engines on a firm mathematical basis. In passing, we cleanly separate the job of the numerical analyst (solving differential equations) from that of the computer scientist (generating execution schemes).

- In late 2010, we presented in 49th Conference on Design and Control in 2010 [11] the use of non standard semantics as a semantical ground for a hybrid synchronous language.

- Since the, we have extended this work in the following directions: 1/ a synchronous Kahn semantics for hybrid programs. Programs are viewed as synchronous ones running on an infinitely fast discrete base clock of the form $BaseClock(\partial) = \{ n\partial | n \in \mathbb{N} \}$, with infinitesimal step $\partial$ and $\mathbb{N}$ the non-standard extension of $\mathbb{N}$. 2/ the definition of a standardization principle that gives sufficient conditions for a hybrid program to be standardizable. Under these conditions, the semantics corresponds to the semantics using super-dense time [44], [40], [42] for hybrid systems defined in [10]. 3/ a large amount of experimentations with Simulink to illustrate some of its pitfalls concerning in particular the treatment of zero-crossing cascades. This work is detailed in a long paper appearing in the *Journal of Computer Science and Systems* [1], in 2011.

12 https://modelica.org/
Starting from a minimal, yet full-featured, Lustre-like synchronous language, we have proposed a conservative extension where data-flow equations can be mixed with ordinary differential equations (ODEs) with possible reset. A type system is proposed to statically distinguish discrete computations from continuous ones and to ensure that signals are used in their proper domains. The extended data-flow language is realized through a source-to-source transformation into a synchronous subset, which can then be compiled using existing tools into routines that are both efficient and bounded in their use of memory. These routines are orchestrated with a single off-the-shelf numerical solver using a simple but precise algorithm which treats causally-related cascades of zero-crossings. We have validated the viability of the approach through experiments with the SUNDIALS library. The basis of this work has been presented at the ACM International Conference on Languages, Compilers, Theory of Embedded Systems, 2011 [3].

This work shows that it is possible to define a language which combines both the expressiveness of synchronous a synchronous language and that of ODEs where continuous solvers are approximated by a black-box solver. The most noticiable result was to recycle several techniques developed for synchronous languages: Kahn semantics, compilation techniques, static analysis. During year 2011, we extended the basic language with with hierarchical automata. This work has been presented at the ACM International Conference on Embedded Software, 2011 [2].

In parallel with these theoretical works, M. Pouzet and T. Bourke have developed during year 2011 a new synchronous language and its compiler. The language, called ZELUS, extends a synchronous language with ODEs. It is first-order, functional and which mixes continuous-time and discrete-time signals. The expressiveness is that of (the first-order subset of) Lucid Synchrone (e.g., type inference and polymorphism, mix of data-flow and hierarchical automata) and ODEs with possible reset. Continuous trajectories are computed by a black-box numerical solver and we made our experiments with SUNDIALS.

6.3. N-Synchronous Languages

The n-synchronous model introduced a way to compose streams which have almost the same clock and can be synchronized through the use of a finite buffer.

We have designed the language Lucy-n to program in this model of computation [4], [48]. This language is similar to the first order synchronous data-flow language Lustre in which a buffer operator is added. A dedicated type system allows to check that programs can be executed in bounded memory and to compute the buffers sizes needed. Technically it is done through the introduction of a subtyping constraint at each bufferization point.

To solve the subtyping constraints we have defined an algorithm that uses an improved version of the state-of-the-art abstraction. We have proved the correctness properties of this new abstraction in Coq.

We also worked on new typing algorithms that do not use clock abstraction and thus allows to model Latency Insensitive Design [5] in Lucy-n [48].

- A. Guatto, together with L. Mandel and M. Pouzet, worked on the code generation for Lucy-n. They investigated two approaches. The first one was to use the schedules and buffer sizes computed by the compiler to generate a classical Lustre program. The second approach was to define a dynamic scheduling protocol similar to the ones used in latency insensitive designs.
- L. Mandel, in collaboration with F. Plateau (Prove&Run), developed a new resolution constraint algorithm for the clocking of Lucy-n programs [4]. Even if the new algorithm is less efficient that the one using abstraction, it has the advantage to be more precise and thus to accept more programs.
- L. Mandel, F. Plateau and M. Pouzet have extended the Lucy-n language with a new operator to be able to model Latency Insensitive Designs. Thanks to the new resolution constraint algorithm, the Lucy-n compiler is able to compute static schedules for such designs [5].

13 https://computation.llnl.gov/casc/sundials/main.html
6.4. Synchronous Circuits

- Followed up on J. Vuillemin’s result that the XOR variant of non-deterministic automata can be efficiently minimized \cite{65}, we explore this newly opened branch of computational automata theory. One contribution is a Decision Diagram for Boolean functions which has minimal dimension: this is appealing for both the verification and synthesis of memory-less circuits. Parts of this recent work were presented at the Boole ANR cooperation \cite{8} and Synchron 2011.

- An extension of Boolean Decision Diagrams to integer representation and operations is given in \cite{66}: we pursue software experimentations with arithmetics on such gigantic (yet sparse) numbers.

- The paper \cite{32} was translated to English from the 1974 original and published in honor of Gilles Kahn.

6.5. Reactive Programming

ReactiveML is an extension of OCaml with synchronous concurrency, based on synchronous parallel composition and broadcast of signals. The goal is to provide a general model of deterministic concurrency inside a general purpose functional language to program reactive systems. It is particularly suited to program discrete simulations, for instance of sensor networks. The current focus of the research is being able to simulate huge systems, composed of millions of agents, by extending the current purely sequential implementation in order to be able to take advantage of multi-core and distributed architectures.

A first experiment consisted in creating a parallel runtime without any modification to the language. As the OCaml language on which ReactiveML is based (the ReactiveML compiler generates OCaml code) does not allow to create parallel programs communicating via shared memory, the new runtime was written in the F# language, part of Microsoft .Net environment. As the language is almost source to source compatible with OCaml, the ReactiveML compiler was left untouched. Several parallel runtimes were written, using traditional task scheduling techniques like work stealing or directly using light task mechanisms available in F#. This experiment demonstrated many speedup opportunities by parallelizing the runtime but also highlighted several problems and limitations of the language. Although this experiment was very useful in understanding the stakes of parallelizing ReactiveML, the performance gap between OCaml and F# (OCaml generates sequential code that is about 10 times slower) makes this version of the runtime of little practical use. We then proposed an extension of ReactiveML called clock domains. It consists in creating local notions of instants that are invisible from the outside. This extension should solve most of the problems raised by the previous experiment and help the parallelization of the language. The sequential runtime was adapted to this extension and a distributed version, using processes communicating via message passing, is currently being developed.

In collaboration with P. Attar (INRIA), F. Boussinot (INRIA) and J.-F. Susini (CNAM), L. Mandel worked on the design of DSL \cite{9}, a script language for the orchestration of concurrent programs. L. Mandel developed in ReactiveML and JoCaml an interpreter for the DSL language.

6.6. Polyhedral Compilation

L.-N. Pouchet presented fundamental advances in the construction and linear optimization of multidimensional affine transformation spaces at the POPL 2011 conference, in collaboration with A. Cohen and colleagues from Paris Sud University, Louisiana State University and Ohio State University.

Our team is actively integrating the polyhedral optimization framework in two production compilers: the Graphite framework in GCC and the Polly framework in LLVM. We are also working towards using the polyhedral framework to target GPGPU and manycore architectures, and to generate aggressively optimized code starting from high level languages. New isl-based version of Graphite and Polly have been contributed, enabling state-of-the-art affine transformations in GCC and LLVM, respectively. Dramatic performance improvements are expected in 2012, impacting the upcoming GCC 4.8 and LLVM 3.1, respectively.

K. Trifunovic, F. Li and A. Cohen, in collaboration with R. Ladelsky from IBM Research Haifa, presented a paper about some of these progresses at the GROW 2011 workshop \cite{7}.
Among the challenges that arise when adapting the polyhedral framework to production compilers, Riyadh Baghdadi has been working on memory-based dependences. Part of it is a practical compiler construction issue, where upstream passes such as the transformation to three-address code and PRE/CSE introduce new scalar variables leading to additional memory-based dependences. The other difficulty is to identify a profitable tradeoff between memory expansion (privatization, renaming) and parallelism. Memory-based dependences not only increase the complexity of the optimization but most importantly, they reduce the degree of freedom available to express effective loop nest transformations, limiting the overall effectiveness of the polyhedral framework. We designed and implemented a technique that solves this problem by allowing a compiler to relax the constraint of memory-based dependences on loop nest transformations and that does not incur the memory footprint overhead of scalar and array expansion. The proposed technique is based on the concept of polyhedral live range interval interference. While previous polyhedral optimization techniques could not achieve any speedups in benchmarks with scalar variables. This technique enabled a speedup of up to $16 \times$ on numerical kernels from the Polybench benchmark suite (on a 24-core machine).

### 6.7. Parallel Data-Flow Programming

A. Pop defended his PhD thesis in September at MINES ParisTech, on a data-flow, streaming extension of OpenMP. Semantical, compilation and runtime system aspects have been covered in depth. Early results were published at HiPEAC 2011 and presented by A. Pop, in collaboration with A. Cohen. Follow-up work include the maturation of the proposed semantics and language extensions, with a thorough implementation and experimentation.

In parallel, F. Li, in collaboration with A. Pop and A. Cohen, explores the automatic compilation of SSA programs into dynamic data-flow parallelism, and the integration of streaming dependences to extend the method to non-scalar data flow. A paper has been published at WIR 2011 (workshop associated to CGO 2011), and a comprehensive, modular compilation method for arbitrary control flow, will be presented at MULTIPROG 2012 (associated with HiPEAC 2012).

Classical compilation techniques, found in Lustre, Scade, Lucid Synchrone, and all the dataflow synchronous languages, generate very efficient sequential code. Thus our main goal is to allow parallel code generation without changing the generation of the sequential code. To this matter, we introduced in the dataflow synchronous setting the famous asynchronous calls bundled with futures, which date back to MultiLisp designed by R. Halstead in the early 1980. It allows to separate the request of a computation from the actual use of its result. This approach has two main advantages. First, the compilation of these asynchronous calls is implemented by a simple wrapper encapsulating the called sequential code. It permits full compatibility with existing generated code. The futures are treated like usual values, so except for the asynchronous calls, we use the known sequential code generation. Second, this asynchronous calls and futures are only annotations and may be fully erased without changing the semantics of the program.

L. Gérard implements this proposition in our Heptagon compiler. The first backend was done in Java, as a proof of concept, using the threads and futures of the Java language. More efficient back-ends are being explored, using OpenMP stream-computing extensions and the TStar data-flow primitives.
6. New Results

6.1. Proof-theoretical investigations


6.1.1. Sequent calculus and Computational duality

Thunks and duality. Guillaume Munch-Maccagnoni investigated a notion dual to the thunks of call-by-value lambda-calculus which allows to simulate call-by-value in call-by-name. He started to investigate how this structure arises in many models of call-by-name lambda-calculus, how it could explain various phenomena such as storage operators, and how it could relate to features of actual programming languages.

Categorical semantics. Guillaume Munch-Maccagnoni started a collaboration with Marcelo Fiore with the aim of understanding structures behind sequent calculus (of system L in particular) and polarisation (as found in Girard’s classical logic). The goal is to draw links with algebraic and/or computing structures in the unifying language of category theory. This work is also in collaboration with Pierre-Louis Curien.

Noam Zeilberger has continued to work with Paul-André Melliès (PPS) on developing a categorical framework for better understanding contexts and inference rules in proof theory and type theory, with the aim of achieving an integration with the theory of side-effects in programming languages. He presented some results from this work in March at the European Workshop on Computational Effects.

Polarised Peano arithmetic. Guillaume Munch-Maccagnoni extended polarised classical logic and polarised classical realizability to predicate calculus and to Peano arithmetic. This decomposes and simplifies technical artefacts found in call-by-name classical realizability, and sheds a new light on witness extraction from proofs of $\Sigma$ formulae.

(Co)Inductive Types in Sequent Calculus. Hugo Herbelin and Jeffrey Sarnat continued their work towards a sequent calculus presentation of a simply-typed fragment of CIC that has inductive and coinductive types, as formalized using recursion operators and a guard condition. Some progress was made on the formalization of the guard condition and normalization proof, but both remained unfinished as of the end of Sarnat’s postdoc in June.

Classical call-by-need and the duality of computation. In a collaboration with Zena Ariola (and especially after a 3-week visit by Alexis Saurin in Oregon early 2011), Zena Ariola, Hugo Herbelin and Alexis Saurin have presented the call-by-need strategy in the framework of the duality of computation, that is a sequent calculus approach to call-by-need, and they extended call-by-need from minimal logic to classical logic which allowed to integrate smoothly control operators, resulting in particular in a call-by-need $\lambda\mu$-calculus. Moreover, the duality principles involved in such a framework unveiled a new calculus, dual to the usual call-by-need but which is distinct from call-by-name, call-by-value as well as the usual call-by-need. These results were presented at TLCA 2011.

Pursuing the previous collaboration, the three previous authors, together with Paul Downen and Keiko Nakata, studied abstract machines for this classical call-by-need calculus.

6.1.2. Linear dependent types

Arnaud Spiwack started investigating dependent types variants of linear sequent calculus based on Curien & Herbelin’s $\mu\tilde{\mu}$. The goal is to study what kind of set theory arises from such a linear type theory when following the tradition of intuitionistic type theory of defining a set as a type equipped with a relation (this construction is also known as setoid). Sets defined in this manner in Coq gives rise to a quasitopos (as proved in Arnaud Spiwack’s PhD thesis) which makes for a reasonable approximation of usual mathematics, however
“linear sets” should be quite different and may support some unorthodox mathematics. To have an appropriate theory of linear sets seems to require a fairly rich linear type theory in particular one that supports so called strong elimination (the type theoretic equivalent to induction). So far, if extending $\mu\tilde{\mu}$ to dependently typed linear logic has been achieved, strong elimination proves to be harder to coin.

Pierre-Marie Pédrot just started his PhD on this same general topic. While linear logic appeared as an operational decomposition of intuitionistic logic, dependent types are conversely an essential enrichment of the latter, as they permit to constructively formalize important parts of mathematics. Even though it is nowadays pervasively thought that we should mix them, actually it does seem that nobody seriously tried hitherto. This subject is quite vast, and both sides may mutually enhance the understanding of one another. From the linear side, linear logic is seriously lacking any satisfying syntax at all; and worse, it seems more prone to solely describe computational behavior instead of truly formalizing mathematics, for its very lack of richer types. From the dependent side, usual dependent type systems are based upon PTS, which, being a plain enrichment of basic lambda-calculus, are intrinsically call-by-name structures. Hence, a linear decomposition inspired from polarization techniques may permit a better analysis of the inner, yet-to-be discovered gears of PTS. One may even hope to include non-intuitionistic effects therein. Furthermore, practical systems used today (Coq, for example) come bundled with additional constructs, such as inductives, whose understanding with respect to models is still highly incomplete. One could expect linear logic to shed a new light upon these issues. This thesis stems from preliminary work of Pierre-Marie Pédrot (during his M2 internship) inspired by models of GoI and other closely related results from Girard, which suggest a natural way to integrate dependency into them.

6.1.3. Proving with side-effects

Axiom of dependent choice. 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 second to turn countable universal quantification into an infinite conjunction of classical proofs evaluated along a call-by-need evaluation strategy so as to extract of them intuitionistic contents that complies to the intuitionistic constraint put on strong elimination of existential quantification. This work has been presented at the TYPES conference.

Memory assignment, forcing and delimited control. Hugo Herbelin investigated how to extend his work on intuitionistically proving Markov’s principle [35] and the work of Danko Ilik on intuitionistically proving the double negation shift (i.e. $\forall x \neg\neg A \rightarrow \neg\neg\forall x A$) [38] to other kind of effects. In particular, memory assignment is related to Cohen’s forcing as emphasized by Krivine [40] and as the observation that Cohen’s translation of formula $P$ into $\forall y \leq x \exists z \leq y P(z)$ is similar to a state-passing-style transformation of type $P$ into $S \rightarrow S \times P$.

Hugo Herbelin then designed a logical formalism with memory assignment that allows to prove in direct-style any statement provable using the forcing method, the same way as logic extended with control operators allows to support direct-style classical reasoning. Thanks to the use of delimiters over “small” formulas similar to the notation of $\Sigma_1$-formulas in arithmetic, the whole framework remains intuitionistic, in the sense that it satisfies the disjunction and existence property.

Two typical applications of proving with side-effects are global-memory proofs of the axiom of countable choice and an enumeration-free proof of Gödel’s completeness theorem.

The main ideas of this research program have been presented at the Geocal-Lac meeting of the GDR IM.

6.1.4. Delimited control

Delimited control and infinitary/stream calculi

During his summer internship, Paul Downen studied with Alexis Saurin some infinitary $\lambda$-calculi and infinitary rewriting and in particular a proposal by Ketema, Blom, Aoto and Simonsen which allows to consider transfinitely deep terms. The proposed calculus presented several defects on which Paul Downen studies focused. Some of these defects were corrected but the work is still on-going.
In a collaboration with Marco Gaboardi and Koji Nakazawa, Alexis Saurin has been studying how to turn the $\Lambda\mu$-calculus into a truly stream-based calculus. This involved enlarging the syntactical category for streams, defining a type system and comparing with other proposals for computing on streams.

Alexis Saurin also developed previous results on $\Lambda\mu$-calculus in a paper currently under final revision for publication in TCS.

**PTS and delimited control**
Following Danvy-Filinski’ simply-typed system for a $\lambda$-calculus with delimited control, Hugo Herbelin and Pierre Boutillier have defined a set of rules for pure type systems with control operator. The work relies on the CPS used to encode them in standard Pure Type Systems and involves extra type annotations. Nevertheless, it seems to be more general than previous attempts to build classical PTS [20]. It has been presented at the workshop TPDC (Theory and Practice of Delimited Continuations) in Novi Sad and an article is under preparation.

### 6.1.5. Interactive Realizability

Thanks to the Curry-Howard correspondence for classical logic, it is possible to extract programs from classical proofs. These programs use control operators as a way to implement backtracking and processes of intelligent learning by trial and error. Unfortunately, such programs are often hard to write, difficult to understand and are very inefficient: every time a program backtracks, it forgets way too much information. This state of thing is due to a poor understanding and control of the backtracking mechanism that interprets classical proofs. In order to write down more efficient programs, it is necessary to describe exactly: a) what the programs learn, b) how the knowledge of programs varies during the execution.

A first step towards this goal is the theory of Interactive Realizability, a semantics for intuitionistic Arithmetic with excluded middle over semi-decidable predicates. It is based on a notion of state, which describes the knowledge of programs coming from a classical proof, and explains how the knowledge evolves during computation.

Federico Aschieri is working in two directions. First he is extending this realizability semantics to second-order intuitionistic Arithmetic with the same excluded middle over semi-decidable predicates. He has also discovered a new state passing style transformation, which allows to implement in System F efficient programs, which backtrack at the right point and do not forget anything when backtracking. He is also investigating an interesting relation with the forcing semantics: it seems that his transformation is a very direct, new constructive formulation of forcing.

Secondly, in collaboration with Berardi, he is also extending Interactive Realizability to first-order Arithmetic, with full excluded middle. This work promises to provide a significantly finer description of the learning mechanism that interprets classical proofs.

### 6.1.6. Substitutions and isomorphisms

Pierre-Louis Curien extended his collaboration with Martin Hofmann (Univ. of Munich) and Richard Garner (MacQuarie University, Sydney), started in 2010, to the point where the picture sought and announced in the report of last year turned out to be a bit less idyllic than expected. Let us recall the question addressed. We wanted to compare precisely two ways of giving a categorical interpretations of Martin-Löf type theory, both overcoming the following mismatch: syntax has exact substitutions, while their categorical interpretation, in terms of pullbacks or fibrations, “implements” substitutions only up to isomorphism. One can then either change the model (strictification) [36], or modify the syntax (by introducing explicit substitutions and more importantly explicit coercions between types that are now only isomorphic) [2]. These approaches turn out to be related through adjunctions in a suitable 2-categorical framework that has a conceptual interest of its own. But these adjunction do not fit entirely together, as we found out early 2011: One is base-dependent, and the other not. So we cannot put them directly aside to get the nice conceptual picture we hoped for. Still, our initial goal of expressing the interpretations in terms of on another can be attained, but this remains to be worked out in detail.
6.1.7. Miscellanea

During his three months visit in Beijing, Pierre-Louis Curien worked on the relations between Rewriting theory and the theory of Gröbner bases and other bases like Janet bases or involutive bases that have been introduced in computer algebra. These comparisons shed some light on the classification of various completion techniques for rewriting systems (a completion turns a rewriting system into an equivalent locally confluent one). This is work in progress.

6.2. Metatheory of Coq and beyond

Participants: Pierre Boutillier, Hugo Herbelin, Yann Régis-Gianas, Jeffrey Sarnat, Vincent Siles, Matthieu Sozeau, Noam Zeilberger.

From the work he has done last year on the Coq termination checker, Pierre Boutillier wrote an article to describe formally the exact new Coq implementation of a structural guard condition that handles commutative cuts.

6.2.1. Calculus of inductive constructions and typed equality

The work of Hugo Herbelin and Vincent Siles on the equivalence of Pure Type Systems with typed or untyped equality has been accepted for publication [11].

6.2.2. Proofs of higher-order programs

Jeffrey Sarnat and Noam Zeilberger continued to investigate the classical program transformations of continuation-passing-style translation and defunctionalisation [46], from the point-of-view of their effect on the termination proofs of higher-order programs. The practical aim of these investigations is to develop a more systematic understanding of termination proofs, which eventually could result in a compiler from proof assistants with higher-order reasoning (such as Coq) to ones with only first-order reasoning (such as Twelf).

6.2.3. Unification

Matthieu Sozeau and Hugo Herbelin worked on improving the unification algorithm of Coq, making it more predictable and resolving important soundness issues (e.g. type-checking, dealing with universes). In collaboration with Beta Ziliani (PhD student at MPI Sarbrucken) and Aleksandar Nanevski (Researcher at IMDEA Madrid), Matthieu Sozeau started a project to formalize (on paper) the improved unification algorithm, taking into account advanced features such as canonical structures and type classes. This will give a detailed view of the system to power users like [33] and improve the system to handle the delicate usage patterns developed in the Mathematical Components team at MSR to scale Coq to large formalizations.

6.3. Coq as a functional programming language

Participants: Stéphane Glondu, Pierre Letouzey, Matthias Puech, Matthieu Sozeau.

6.3.1. Type-classes and libraries

Matthieu Sozeau has worked with members of the Foundations team at Nijmegen on enhancing the implementation of type-classes to suit the needs of the development of the MathClasses formalization of abstract algebra (part of the ForMath EU project). This gave rise to an experimental implementation of forward reasoning for instance resolution. Concurrently, he started a collaboration with Jael Kriener (Univ. of Canterbury, who visited for a week in November) to adapt techniques from logic programming to the system, including a determinacy analysis that would give better control of the system when building large hierarchies of structures (as in the MathClasses library).

Pierre Castéran from INRIA Bordeaux and Matthieu Sozeau are developing a tutorial on the use of type-classes that will be the basis of an invited tutorial given by M. Sozeau at the JFLA conference in February 2012. It will be published as part of the new version of the Coq’Art book.
6.3.2. Equations

Matthieu Sozeau has continued the development of the Equations package to build and reason on dependently-typed programs. He reworked the internals of the tool for a more efficient implementation and extended it to handle recursion on arbitrary inductive families. He developed a sizable example of the use of the tool for proving the metatheory of the LF system. The plugin will be relased with the upcoming 8.4 version of Coq. A journal article presenting the tool and its usage is in preparation.

6.3.3. Recursive definitions in proof assistants

Together with Ana Bove and Alexander Krauss, Matthieu Sozeau has written a survey article on tools and methods to build and reason on recursive functions in proof assistants based on type theory. The survey compares the relative strenghts and weaknesses of various formalization methods available in constructive type theories like Agda or Coq and classical systems like HOL. The article is still under review.

6.3.4. Dependent pattern-matching

Hugo Herbelin and Pierre Boutillier worked on a new simulation of Agda’s style dependent pattern-matching [24] in the Calculus of Inductive Constructions which, on the contrary of Goguen et al’s simulation [32], does not rely on any explicit notion of dependent equality. The simulation relies on a systematic re-generalization of a given matching over the next pattern to matching and over a notion of matching narrowed by type constraints formalized by Pierre Boutillier [14]. The former itself relies on Pierre Boutillier’s master thesis on making the guard condition traversing blocked commutative cuts. The simulation has been implemented in Coq 8.4 but without narrowing on subpatterns yet.

6.3.5. Modularized arithmetical libraries

Pierre Letouzey has integrated in Coq a deep reform of some parts of its Standard Library, mainly the Numbers library of generic / efficient arithmetic. This reform is part of the the version Coq 8.4. The idea is to take advantage of recent improvements of the Coq system in terms of modularity (Type Classes by Sozeau and better Modules by Soubiran) for providing more uniformity in the functions and properties about integers provided in the Standard Library. We now have a base of functions and lemmas which is statically guaranteed to be coherent from one numerical datatype representation to the other, allowing the user to choose more easily between these representations, according to the user’s need in term of simplicity or efficiency. These modernized libraries are also one of the first large-scale experimentations with many recent features of modules and type-classes, and have helped maturing them and established new usage guidelines. The transition from previous versions of these libraries should be relatively transparent thanks to a compatibility layer.

6.3.6. Certified extraction

Stéphane Glondu continued his work on extraction in the Coq-in-Coq formalisation. He proved that reductions in the source language can be simulated in the target language (the converse had been proven two years before). This proof highlighted a critical bug in the actual implementation of Coq that has been solved by Pierre Letouzey.

The formalised extraction is only a step of the actual implementation of Coq: replacing logical subterms by an inert constant. Stéphane Glondu considers this work done, even though some parts have been admitted. He worked on how the formalisation could be redesigned in order to avoid currently admitted lemmas and to allow a better integration with the existing Coq system.

6.3.7. Incrementality in proof languages

Matthias Puech and Yann Régis-Gianas worked on incremental type checking. This preliminary work will be presented during a contributed talk at TLDI 2012 [15]. It sets the grounds of an incremental proof development and checking system, by means of a representation language for repositories of proofs and proof changes.
The traditional interaction with a proof-checker is a batch process. Coq (among others) refines this model by providing a read-eval print loop with a global undo system, implemented in an ad-hoc way. A more general approach to incrementality is being developed by means of a finer-grained analysis of dependencies. The approach developed is adaptable to any typed formal language: the language is specified in a meta-language close to the Edinburgh Logical Framework, in which subsequent versions of a development can be type-checked incrementally. Applications of this framework are: proof language for proof assistants, integrated development environments for proof or programming languages, typed version control systems.

6.3.8. Proofs of programs in Coq

As part of the CerCo european project, in collaboration with Roberto Amadio (PPS, Paris Diderot University), Nicolas Ayache and Yann Régis-Gianas maintained a prototype compiler for a large subset of the C language whose specificity is to annotate input source programs with information about the worst-case execution cost of their machine code. Yann Régis-Gianas started a mechanized version of the proof technique used to prove the correctness of such an annotating compiler.

Yann Régis-Gianas developed another compiler for Core ML that uses a generalization of CerCo technique in order to obtain certified worst case execution time bounds on functional programs. This compiler produces proof obligations in Coq. The corresponding paper will be published in January 2012 in the proceedings of FOPARA 2011. It is available as a technical report [16].
6. New Results

6.1. Real Solving Polynomial Systems

In [20], we describe an algorithm (VQE) for a variant of the real quantifier elimination problem (QE). The variant problem requires the input to satisfy a certain extra condition, and allows the output to be almost equivalent to the input. The motivation/rationale for studying such a variant QE problem is that many quantified formulas arising in applications do satisfy the extra conditions. Furthermore, in most applications, it is sufficient that the output formula is almost equivalent to the input formula. The main idea underlying the algorithm is to substitute the repeated projection step of CAD by a single projection without carrying out a parametric existential decision over the reals. We find that the algorithm can tackle important and challenging problems, such as numerical stability analysis of the widely-used MacCormack’s scheme. The problem has been practically out of reach for standard QE algorithms in spite of many attempts to tackle it. However, the current implementation of VQE can solve it in about 12 hours. This paper extends the results reported at the conference ISSAC 2009.

We also focused on the interaction of real solving polynomial system with global optimization. Let \( f \in \mathbb{Q}[X_1, \ldots, X_n] \) of degree \( D \). Algorithms for solving the unconstrained global optimization problem \( f^* = \inf_{x \in \mathbb{R}^n} f(x) \) are of first importance since this problem appears frequently in numerous applications in engineering sciences. This can be tackled by either designing appropriate quantifier elimination algorithms or by certifyng lower bounds on \( f^* \) by means of sums of squares decompositions but there is no efficient algorithm for deciding if \( f^* \) is a minimum. The paper [41] is dedicated to this important problem. We design an algorithm that decides if \( f^* \) is reached over \( \mathbb{R}^n \) and computes a point \( x^* \in \mathbb{R}^n \) such that \( f(x^*) = f^* \) if such a point exists. If \( L \) is the length of a straight-line program evaluating \( f \), a probabilistic version of the algorithm runs in time \( \tilde{O}(n^2(L + n^2)(D(D - 1)^{n-1})^2) \). Experiments show its practical efficiency.

Global optimization problems can also be tackled by computing algebraic certificates of positivity through sums of squares decompositions. Let \( f_1, \ldots, f_p \) and \( f \) be polynomials in \( \mathbb{Q}[X_1, \ldots, X_n] \) and let \( V = V(f_1, \ldots, f_p) \subset \mathbb{C}^n \) be the algebraic variety defined by \( f_1 = \cdots = f_p = 0 \) whose dimension is denoted by \( d \). Assume in the sequel that the ideal \( \langle f_1, \ldots, f_p \rangle \) is radical and equidimensional and that \( V \) is smooth. In [18], up to a generic linear change of variables, we construct families of polynomials \( M_0, \ldots, M_d \) in \( \mathbb{Q}[X_1, \ldots, X_n] \) such that \( f(x) \geq 0 \) for all \( x \in V \cap \mathbb{R}^n \) if and only if \( f \) can written as a sum of squares of polynomials in \( \mathbb{R}[X_1, \ldots, X_n] \) modulo \( \langle M_i \rangle \) for \( 0 \leq i \leq d \). Such an algebraic certificate of positivity is simpler than the more general Positivstellensatz in Real Algebra. It can be used to certify lower bounds on \( f^* = \inf_{x \in V \cap \mathbb{R}^n} f(x) \). Also, computing a numerical approximation of such certificates of positivity can be reformulated as a semidefinite program which can be solved efficiently. We provide numerical experiments showing the effectiveness of our approach.

In [25], we consider the problem of constructing roadmaps of real algebraic sets. This problem was introduced by Canny to answer connectivity questions and solve motion planning problems. Given \( s \) polynomial equations with rational coefficients, of degree \( D \) in \( n \) variables, Canny’s algorithm has a Monte Carlo cost of \( s^n \log(s)D^{O(n^2)} \) operations in \( \mathbb{Q} \); a deterministic version runs in time \( s^n \log(s)D^{O(n^2)} \). A subsequent improvement was due to Basu, Pollack and Roy, with an algorithm of deterministic cost \( s^{d+1}D^{O(n^2)} \) for the more general problem of computing roadmaps of a semi-algebraic set (\( d \leq n \) is the dimension of an associated object). We give a probabilistic algorithm of complexity \( (nD)^{O(n^{1.5})} \) for the problem of computing a roadmap of a closed and bounded hypersurface \( V \) of degree \( D \) in \( n \) variables, with a finite number of singular points. Even under these extra assumptions, no previous algorithm featured a cost better than \( D^{O(n^2)} \).
6.2. Zero dimensional Solve

Let \( I \subset \mathbb{K}[x_1, \ldots, x_n] \) be a 0-dimensional ideal of degree \( D \) where \( \mathbb{K} \) is a field. It is well-known that obtaining efficient algorithms for change of ordering of Gröbner bases of \( I \) is crucial in polynomial system solving. Through the algorithm \( \text{FGLM} \), this task is classically tackled by linear algebra operations in \( \mathbb{K}[x_1, \ldots, x_n]/I \).

With recent progress on Gröbner bases computations, this step turns out to be the bottleneck of the whole solving process.

In [38], we present an efficient algorithm that takes advantage of the sparsity structure of multiplication matrices appearing during the change of ordering. This sparsity structure arises even when the input polynomial system defining \( I \) is dense. As a by-product, we obtain an implementation which is able to manipulate 0-dimensional ideals over a prime field of degree greater than 30000. It outperforms the \( \text{Magma/Singular/FGb} \) implementations of \( \text{FGLM} \).

In [38], we investigate the particular but important shape position case. The obtained algorithm performs the change of ordering within a complexity \( O(D(N_1 + n \log(D))) \), where \( N_1 \) is the number of nonzero entries of a multiplication matrix (the density of the matrix). This almost matches the complexity of computing the minimal polynomial of one multiplication matrix. Then, we address the general case and give corresponding complexity results. Our algorithm is dynamic in the sense that it selects automatically which strategy to use depending on the input. Its key ingredients are the Wiedemann algorithm to handle 1-dimensional linear recurrence (for the shape position case), and the Berlekamp–Massey–Sakata algorithm from Coding Theory to handle multi-dimensional linearly recurring sequences in the general case.

6.3. Solving structured systems

Solving multihomogeneous systems, as a wide range of structured algebraic systems occurring frequently in practical problems, is of first importance. Experimentally, solving these systems with Gröbner bases algorithms seems to be easier than solving homogeneous systems of the same degree. Nevertheless, the reasons of this behaviour are not clear. In [16], we focus on bilinear systems (i.e. bihomogeneous systems where all equations have bidegree \( (1, 1) \)). Our goal is to provide a theoretical explanation of the aforementioned experimental behaviour and to propose new techniques to speed up the Gröbner basis computations by using the multihomogeneous structure of those systems. The contributions are theoretical and practical.

First, we adapt the classical \( F_5 \) criterion to avoid reductions to zero which occur when the input is a set of bilinear polynomials. We also prove an explicit form of the Hilbert series of bihomogeneous ideals generated by generic bilinear polynomials and give a new upper bound on the degree of regularity of generic affine bilinear systems. This leads to new complexity bounds for solving bilinear systems. We propose also a variant of the \( F_5 \) Algorithm dedicated to multihomogeneous systems which exploits a structural property of the Macaulay matrix which occurs on such inputs. Experimental results show that this variant requires less time and memory than the classical homogeneous \( F_5 \) Algorithm.

Lasty, we investigate the complexity of computing a Gröbner basis for the grevlex ordering of a generic 0-dimensional affine bilinear system over \( \mathbb{K}[x_1, \ldots, x_n, y_1, \ldots, y_m] \). In particular, we show that this complexity is upper bounded by \( O\left(\left(\min(n_x + 1, n_y + 1)\right)^{N} \right) \), which is polynomial in \( n_x + n_y \) (i.e. the number of unknowns) when \( \min(n_x, n_y) \) is constant.

6.4. Structured Systems and Applications to Cryptanalysis

The Goppa Code Distinguishing (GCD) problem consists in distinguishing the matrix of a Goppa code from a random matrix. Up to now, it is widely believed that the GCD problem is a hard decisional problem. In [36], we present the first technique allowing to distinguish alternant and Goppa codes over any field. Our technique can solve the GCD problem in polynomial-time provided that the codes have rates sufficiently large. The key ingredient is an algebraic characterization of the key-recovery problem which reduces to the solving of a system of bi-homogeneous polynomial equations. The idea is to consider the dimension of the solution space of a linearized system deduced from the algebraic system describing the key-recovery. It turns
out that experimentally this dimension depends on the type of code. Explicit formulas derived from extensive experimentations for the value of the dimension are provided for “generic” random, alternant, and Goppa code over any alphabet. Finally, we give explanations of these formulas in the case of random codes, alternant codes over any field and binary Goppa codes.

6.5. Algebraic Cryptanalysis

The Isomorphism of Polynomials (IP) is one of the most fundamental problems in multivariate public key cryptography (MPKC). In In [23], we introduce a new framework to study the counting problem associated to IP. Namely, we present tools of finite geometry allowing to investigate the counting problem associated to IP. Precisely, we focus on enumerating or estimating the number of isomorphism equivalence classes of homogeneous quadratic polynomial systems. These problems are equivalent to finding the scale of the key space of a multivariate cryptosystem and the total number of different multivariate cryptographic schemes respectively, which might impact the security and the potential capability of MPKC. We also consider their applications in the analysis of a specific multivariate public key cryptosystem. Our results not only answer how many cryptographic schemes can be derived from monomials and how big the key space is for a fixed scheme, but also show that quite many HFE cryptosystems are equivalent to a Matsumoto-Imai scheme.

In [34], we present a practical cryptanalysis of the Identification Scheme proposed by Patarin at Crypto 1996. This scheme relies on the hardness of the Isomorphism of Polynomial with One Secret (IP1S), and enjoys shorter key than many other schemes based on the hardness of a combinatorial problem (as opposed to number theoretic problems). We present two new deterministic algorithms to attack the IP1S problem, and we rigorously analyze their complexity and success probability. We show that they can solve a constant fraction of all the instances of degree two in polynomial time.

In [33], we investigate the security of a generalization of HFE (multivariate and odd-characteristic variants). We propose an improved version of the basic Kipnis-Shamir key recovery attack against HFE. We then generalize the Kipnis-Shamir attack to Multi-HFE. The attack reduces to solve a MinRank problem directly on the public key. This leads to an improvement of a factor corresponding to the square of the degree of the extension field. We used recent results on MinRank to show that our attack is polynomial in the degree of the extension field. It appears that multi-HFE is less secure than original HFE for equal-sized keys. Finally, adaptations of our attack overcome several variants (i.e. minus modifier and embedding). As a proof of concept, we have practically broken the most conservative parameters given by Chen, Chen, Ding, Werner and Yang in 9 days for 256 bits security. All in all, our results give a more precise picture on the (in)security of several variants of HFE proposed these last years.

In [31], we initiate the formal treatment of cryptographic constructions (“Polly Cracker”) based on the hardness of computing remainders modulo an ideal. We start by formalising and studying the relation between the ideal remainder problem and the problem of computing a Gröbner basis. We show both positive and negative results. On the negative side, we define a symmetric Polly Cracker encryption scheme and prove that this scheme only achieves bounded CPA security under the hardness of the IR problem. Furthermore, we show that a large class of algebraic transformations cannot convert this scheme to a fully secure Polly Cracker-style scheme. On the positive side, we formalise noisy variants of the ideal related problems. These problems can be seen as natural generalisations of the LWE problem and the approximate GCD problem over polynomial rings. After formalising and justifying the hardness of the noisy assumptions we show that noisy encoding of messages results in a fully IND-CPA secure somewhat homomorphic encryption scheme. Together with a standard symmetric-to-asymmetric transformation for additively homomorphic schemes, we provide a positive answer to the long standing open problem of constructing a secure Polly Cracker-style cryptosystem reducible to the hardness of solving a random system of equations. Indeed, our results go beyond that by also providing a new family of somewhat homomorphic encryption schemes based on new, but natural, hard problems. Our results also imply that Regev’s LWE-based public-key encryption scheme is (somewhat) multiplicatively homomorphic for appropriate choices of parameters.
6.6. Computer Algebra and Algorithmic Number Theory

The Elliptic Curve Discrete Logarithm Problem (ECDLP) has become the most attractive alternative to factoring for public key cryptography. Whereas subexponential factoring algorithms exist, solving the ECDLP in general can only be done in exponential time. Provided that a certain heuristic assumption holds, we present in [39] an index calculus algorithm solving ECDLP over any binary field $\mathbb{F}_q$ in time $O(2^{e \cdot n^{2/3} \log n})$, where $e$ is a small constant. Our algorithm follows the index calculus method that was first introduced by Semaev and later developed by Gaudry and Diem. In particular, the main step consists in decomposing points of the curve with respect to an appropriately chosen factor basis. This part can be nicely reformulated as a purely algebraic problem consisting in finding solutions to a multivariate polynomial $f(x_1, \ldots, x_m) = 0$ such that all the variables $x_i$ belong to some vector subspace of $\mathbb{F}_q/\mathbb{F}_p$. We solve this problem by means of Gröbner basis techniques and analyse its complexity using the multihomogeneous structure of the equations. Even, if this paper is essentially theoretical and is not aiming for practical attacks, the new ideas developed here could be used to have practical attacks in the future. This of course represents a challenging open problem.
5. New Results

5.1. Symmetric cryptosystems

Participants: Céline Blondeau, Christina Boura, Baudoin Collard, Anne Canteaut, Pascale Charpin, Stéphane Jacob, Gohar Kyureghyan.

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 as 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 restricting implementation requirements are crucial when designing secure symmetric primitives and they might be at the origin of some weaknesses. Actually, these constraints seem quite incompatible with the rather complex mathematical tools needed for constructing a provably secure system.

The specificity of our research work is that it considers all aspects of the field, from the practical ones (new attacks, concrete specifications of new systems) to the most theoretical ones (study of the algebraic structure of underlying mathematical objects, definition of optimal objects). But, our purpose is to study these aspects not separately but as several sides of the same domain. Our approach mainly relies on the idea that, in order to guarantee a provable resistance to the known attacks and to achieve extremely good performance, a symmetric cipher must use very particular building blocks, whose algebraic structures may introduce unintended weaknesses. Our research work captures this conflict for all families of symmetric ciphers. It includes new attacks and the search for new building blocks which ensure both a high resistance to the known attacks and a low implementation cost. This work, which combines cryptanalysis and the theoretical study of discrete mathematical objects, is essential to progress in the formal analysis of the security of symmetric systems.

In this context, the very important challenges are the designs of low-cost ciphers and of secure hash functions. Most teams in the research community are actually working on the design and on the analysis (cryptanalysis and optimization of the performance) of such primitives.

5.1.1. Hash functions.

Following the recent attacks against almost all existing hash functions (MD5, SHA-0, SHA-1...), we have initiated a research work in this area, especially within the Saphir-2 ANR project and with several PhD theses. Our work on hash functions is two-fold: we have designed two new hash functions, named FSB and Shabal, which have been submitted to the SHA-3 competition, and we have investigated the security of several hash functions, including the previous standards (SHA-0, SHA-1...) and some other SHA-3 candidates.

Recent results:

• study of the algebraic properties of the recent hash function proposals, including the SHA-3 candidates Keccak and Luffa. This work includes a theoretical study of the algebraic degree of iterated functions composed of parallel applications of a smaller function [24].

• Upper bounds on the degree of an iterated permutation from the degree of the inverse of the inner transformation; this result has been applied both to hash functions and to block ciphers [31], [44].

5.1.2. Stream ciphers.

Our research work on stream ciphers is a long-term work which has been developed within the 4-year ANR RAPIDE project. It includes an important cryptanalytic effort on stream ciphers.
Recent results:

- Evaluation of the bias of parity-check relations in the context of cryptanalysis of combination generators with constituent devices which generate period sequences [13].
- Cryptanalysis of the recent stream cipher proposal Armadillo [21].

5.1.3. Block ciphers.

Even if the security of the current block cipher standard, AES, is not threaten 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.

Recent results:

- Differential cryptanalysis with multiple differentials, multiple differential cryptanalysis on the lightweight block cipher Present [23].
- Use of tools from error correcting theory in linear cryptanalysis [36].
- Determination of the data complexity (i.e., of the required number of plaintexts-ciphertexts) and of the success probability of all statistical attacks against block ciphers [12].

5.1.4. Cryptographic properties and construction of appropriate building blocks.

The construction of building blocks which guarantee a high resistance to 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 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.

Recent results:

- Study of the properties of the family of power functions with exponents \(2^t - 1\). This family notably includes the cube function \(x^3\) and the inverse function over a finite field with characteristic 2. In this work, the whole Walsh spectrum of \(x^7\) is determined [11].
- Construction and study of the properties of new families of permutation polynomials over the field with \(2^n\) elements; study of permutations with a linear structure: [14].
- Study of the algebraic properties (e.g. the algebraic degree) of the inverses of APN power permutations [47].

5.2. Code-based cryptography

Participants: Matthieu Finiasz, Grégory Landais, Rafael Misoczki, Ayoub Otmani, 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 \((\mathbb{Z}/n\mathbb{Z})\) 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 scheme).

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,
- address new functionalities, like hashing or symmetric encryption.

Recent results:

- A distinguishing attack on high rate Goppa codes [25]. This results does not lead to an attack on any code based cryptosystem, but, in some particular cases, it invalidates the security reduction. It was conjectured that there was no such distinguishers.
- A new class of codes for McEliece type cryptosystems offering more versatility [22]
- A generic attack on one-time signature based on codes (KKS type) [27].
- A improvement of generic decoding techniques when addressing multiple targets [28].

### 5.3. Error-correcting codes and applications

**Participants:** Mamdouh Abbara, Matthieu Finiasz, Vincent Herbert, Denise Maurice, Nicolas Sendrier, Jean-Pierre Tillich.

Decoding algorithms are extensively used for cryptanlyses. For instance, a classical cryptanalysis of the stream ciphers which rely on linear feedback shift register filtered by a Boolean function models the attacked cipher as the result of the transmission of a linear function through a very highly noisy channel. Then, removing the noise amounts to decoding a certain linear code. This code is highly structured, and one of the most efficient methods to decode it exploits the fact that it has low density parity-check equations, and thus can be decoded as a low-density parity-check code, with iterative algorithms. Furthermore, the problem of finding good approximations of ciphers amounts to a decoding problem of the first order Reed-Muller code. Local decoding is then used in this context, and enables various attacks, such as correlation attacks or linear cryptanalysis.

Besides the cryptographic applications of decoding algorithms, we also investigate two new application domains for decoding algorithms: reverse engineering of communication systems, and quantum error correcting codes for which we have shown that some of them can be decoded successfully with iterative decoding algorithms.
5.3.1. Algebraic error-correcting codes.

Finding lower bounds on the minimum distance of cyclic codes is an old and difficult problem. Cyclic codes with three zeroes correct at most three errors, that is have minimum distance at most 7. It is an interesting question to determine which cyclic codes with three zeroes have minimum distance 7. Vincent Herbert revisit this problem by using an algorithm due to Shaub. Some classification questions are addressed about three error correcting cyclic codes and some new results involving intensive computer search have been obtained [10], [26].

5.3.2. Quantum codes.

The knowledge we have acquired in iterative decoding techniques has also led to study whether or not the very same techniques could also be used to decode quantum codes. Part of the old ACI project “RQ” in which we were involved and the new ANR project “COCQ” are about this topic. It is worth noticing that 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:

- a construction of a family of quantum turbo-codes with excellent error reducing performance under iterative decoding and this even for very noisy channels [29];
- a proof that this family has unbounded minimum distance [20].

5.3.3. Reverse engineering of communication systems.

To evaluate the quality of a cryptographic algorithm, it is usually assumed that its specifications are public, as, in accordance with Kerckhoffs principle\(^2\), 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, is observed out of a noisy channel. To access the information, the whole communication system has first to be disassembled and every constituent reconstructed. Our activity within this domain, whose first aim is to establish the scientific and technical foundations of a discipline which does not exist yet at an academic level, has been supported by some industrial contracts driven by the DGA.

---

\(^2\)Kerckhoffs stated that principle in a paper entitled _La Cryptographie militaire_, published in 1883.
6. New Results

6.1. Robustness and Tolerance

6.1.1. Cubic B-spline approximation by curve unclamping
Participants: Xiao-Diao Chen, Wei-Hyn Ma, Jean-Claude Paul.

A new approach for cubic B-spline curve approximation is presented. The method produces an approximation cubic B-spline curve tangent to a given curve at a set of selected positions, called tangent points, in a piecewise manner starting from a seed segment. A heuristic method is provided to select the tangent points. The first segment of the approximation cubic B-spline curve can be obtained using an inner point interpolation method, least-squares method or geometric Hermite method as a seed segment. The approximation curve is further extended to other tangent points one by one by curve unclamping. New tangent points can also be added, if necessary, by using the concept of the minimum shape deformation angle of an inner point for better approximation. Numerical examples show that the new method is effective in approximating a given curve and is efficient in computation [ 2 ].

6.1.2. Computing the Hausdorff distance between two B-spline curves
Participants: Xiao-Diao Chen, Wei-Hyn Ma, Gang Xu, Jean-Claude Paul.

This paper presents a geometric pruning method for computing the Hausdorff distance between two B-spline curves. It presents a heuristic method for obtaining the one-sided Hausdorff distance in some interval as a lower bound of the Hausdorff distance, which is also possibly the exact Hausdorff distance. Then, an estimation of the upper bound of the Hausdorff distance in an sub-interval is given, which is used to eliminate the sub-intervals whose upper bounds are smaller than the present lower bound. The conditions whether the Hausdorff distance occurs at an end point of the two curves are also provided. These conditions are used to turn the Hausdorff distance computation problem between two curves into a minimum or maximum distance computation problem between a point and a curve, which can be solved well. A pruning technique based on several other elimination criteria is utilized to improve the efficiency of the new method. Numerical examples illustrate the efficiency and the robustness of the new method [ 3 ].

6.1.3. Surface area estimation of digitized 3D objects using quasi-Monte Carlo methods
Participants: Yu-Shen Liu, Jing Yi, Hu Zhang, Guo-Qin Zheng, Jean-Claude Paul.

A novel and efficient quasi-Monte Carlo method for estimating the surface area of digitized 3D objects in the volumetric representation is presented. It operates directly on the original digitized objects without any surface reconstruction procedure. Based on the Cauchy-Crofton formula from integral geometry, the method estimates the surface area of a volumetric object by counting the number of intersection points between the object’s boundary surface and a set of uniformly distributed lines generated with low-discrepancy sequences. Using a clustering technique, we also propose an effective algorithm for computing the intersection of a line with the boundary surface of volumetric objects. A number of digitized objects are used to evaluate the performance of the new method for surface area measurement [ 5 ].

6.1.4. Reverse Engineering for NC Machining Simulation
Participants: Nabil Anwer, Yi-Jun Yang, Haibi Zhao, Olivier Coma, Jean-Claude Paul.
Reverse engineering for NC Machining simulation is becoming an important component of NC simulation and verification. Design engineers need more accurate and complete CAD model of the simulated machined part for finite element analysis or parametric feature-based modeling for design modification or update. The as-cut or in process geometry should be correctly accessed in the CAD/CAM environment at any stage of the machining process. Few commercial software are addressing the reverse engineering issue and provide robust solutions. Until now, in process CAD models for NC simulation have been created with many drawbacks and inaccurate methods are proposed. Reverse engineering for NC machining simulation based on polyhedral in-process geometry is addressed. Two complementary approaches are presented here. An enriched representation embedded in “Spring technologies Reverse engineering” or SRE file format enables to convert the polyhedral model to STEP file and a discrete shape recognition and segmentation approach provides a promising issue thanks to discrete differential geometry [43].

6.1.5. Projection of curves on B-spline surfaces using quadratic reparameterization  
**Participants:** Yi-Jun Yang, Wei Zeng, Hui Zhang, Jun-Hai Yong, Jean-Claude Paul.

Curves on surfaces play an important role in computer aided geometric design. In this paper, we present a hyperbola approximation method based on the quadratic reparameterization of Bezier surfaces, which generates reasonable low degree curves lying completely on the surfaces by using iso-parameter curves of the reparameterized surfaces. The Hausdorff distance between the projected curve and the original curve is controlled under the user-specified distance tolerance. The projected curve is $G^1$ continuous, where $T$ is the user-specified angle tolerance. Examples are given to show the performance of our algorithm [64].

6.1.6. A kind of parametric transform for trimmed surfaces  
**Participants:** Sheng Yang, Jun-Hai Yong.

An approach for a kind of parametric transform is presented for trimmed parametric surfaces. Firstly, the characteristics of trimmed surface before and after parametric transform are evaluated. Then, an algorithm is proposed to adjust the geometric and topological data of a trimmed surface in order to achieve consistency. At last, the trimmed sphere surface is taken as example to further illustrate the algorithm [63].

6.1.7. A cell-based algorithm for evaluating directional distances in GIS  
**Participants:** Sheng Yang, Jun-Hai Yong, Jia-Guang Sun, He-Jin Gu, Jean-Claude Paul.

Directional distance is commonly used in geographical information systems as a measure of openness. In previous works, the sweep line method and the interval tree method have been employed to evaluate the directional distances on vector maps. Both methods require rotating original maps and study points in every direction of interest. In this article, we propose a cell-based algorithm that pre-processes a map only once; that is, it subdivides the map into a group of uniform-sized cells and records each borderline of the map into the cells traversed by its corresponding line segment. Based on the pre-processing result, the neighbouring borderlines of a study point can be directly obtained through the neighbouring cells of the point, and the borderlines in a definite direction can be simply acquired through the cells traversed by the half line as well. As a result, the processing step does not need to enumerate all the borderlines of the map when determining whether a point is on a borderline or finding the nearest intersection between a half line and the borderlines. Furthermore, we implement the algorithm for determining fetch length in coastal environment. Once the pre-processing is done, the algorithm can work in a complex archipelago environment such as to calculate the fetch lengths in multiple directions, to determine the inclusion property of a point, and to deal with the singularity of a study point on a borderline [61].

6.1.8. A point-in-polygon method based on a quasi-closest point  
**Participants:** Sheng Yang, Jun-Hai Yong, Jia-Guang Sun, He-Jin Gu, Jean-Claude Paul.
This paper presents a numerically stable solution to a point-in-polygon problem by combining the orientation method and the uniform subdivision technique. We define first a quasi-closest point that can be locally found through the uniform subdivision cells, and then we provide the criteria for determining whether a point lies inside a polygon according to the quasi-closest point. For a large number of points to be tested against the same polygon, the criteria are employed to determine the inclusion property of an empty cell as well as a test point. The experimental tests show that the new method resolves the singularity of a test point on an edge without loss of efficiency. The GIS case study also demonstrates the capability of the method to identify which region contains a test point in a map [62].

6.1.9. Using diffusion distances for flexible molecular shape comparison

**Participants:** Yu-Shen Liu, Qi Li, Guo-Qin Zheng, Karthik Ramani, William Benjamin.

Background: Many molecules are flexible and undergo significant shape deformation as part of their function, and yet most existing molecular shape comparison (MSC) methods treat them as rigid bodies, which may lead to incorrect shape recognition. In this paper, we present a new shape descriptor, named Diffusion Distance Shape Descriptor (DDSD), for comparing 3D shapes of flexible molecules. The diffusion distance in our work is considered as an average length of paths connecting two landmark points on the molecular shape in a sense of inner distances. The diffusion distance is robust to flexible shape deformation, in particular to topological changes, and it reflects well the molecular structure and deformation without explicit decomposition. Our DDSD is stored as a histogram which is a probability distribution of diffusion distances between all sample point pairs on the molecular surface. Finally, the problem of flexible MSC is reduced to comparison of DDSD histograms [50].

6.1.10. A fast sweeping method for computing geodesics on triangular manifolds

**Participants:** Song-Gang Xu, Yun-Xiang Zhang.

A wide range of applications in computer intelligence and computer graphics require computing geodesics accurately and efficiently. The fast marching method (FMM) is widely used to solve this problem, of which the complexity is $O(N \log N)$, where $N$ is the total number of nodes on the manifold. A fast sweeping method (FSM) is proposed and applied on arbitrary triangular manifolds of which the complexity is reduced to $O(N)$. By traversing the undigraph, four orderings are built to produce two groups of interfering waves, which cover all directions of characteristics. The correctness of this method is proved by analyzing the coverage of characteristics. The convergence and error estimation are also presented.

6.1.11. A torus patch approximation approach for point projection on implicit surface

**Participants:** Xiao-Ming Liu, Lei Yang.

Point projection on an implicit surface is essential for the geometric modeling and graphics applications of it. This paper presents a method for computing the principle curvatures and principle directions of an implicit surface. Using the principle curvatures and principle directions, we construct a torus patch to approximate the implicit surface locally. The torus patch is second order osculating to the implicit surface. By taking advantage of the approximation torus patch, this paper develops a second order geometric iterative algorithm for point projection on the implicit surface. Experiments illustrate the efficiency and less dependency on initial values of our algorithm [52].

6.1.12. Shape similarity assessment approach for CAD models based on graph edit distance

**Participants:** Bin Wang, Dong Li.

This paper proposes a new shape similarity assessment approach for CAD models in Boundary Representation (Brep) based on graph edit distance. A suboptimal computational procedure is performed to find the best alignment between local structures sets of attributed graphs derived from models. Assuming that only a minority of local structures characterize the functionality, we figure out the weight of every local structure in the query model through a training phase, and then evaluate the similarity between two models by calculating the weighted graph edit distance of corresponding attributed graphs. Experiment results show that our method provides solid retrieval performance on a real-world CAD model database [56].
6.1.13. **Reconstructing 3D objects from 2D sectional views of engineering drawings using volume-based method**

**Participants:** Yamei Wen, Hui Zhang, Zhongmian Yu, Jia-Guang Sun, Jean-Claude Paul.

Sectional views are widely used in engineering practice due to their clear and concise expression. However, it is difficult for computers to understand because of the large numbers of omitted entities and their diversified representations. This paper aims at reconstructing 3D models from 2D sectional views by improving the traditional volume based method. First, we present a two-stage loop searching algorithm to extract desired loops from sectional views. Then, sub-objects are identified by the hint-based feature identification algorithm with an intuitive loop-matching criterion. After that, a model-directed algorithm is proposed to guide the generation of sub-objects which are assembled together to form the final objects. The algorithm can handle full sections, partial sections and offset sections, as well as orthographic views. Multiple sectional views are supported in our algorithm. Moreover, the domain of objects is extended to inclined quadric surfaces and intersecting quadric surfaces with higher order curves. Experiment results show its practicability [60].

6.1.14. **Conditions for coincidence of two cubic Beziér curves**

**Participants:** Wen-Ke Wang, Hui Zhang, Xiao-Ming Liu, Jean-Claude Paul.

This paper presents a necessary and sufficient condition to judge whether two cubic Beziér curves are coincident. For two cubic Beziér curves whose control points are not collinear, they are coincident if and only if their corresponding control points are coincident or one curve is the reversal of another curve. However, this is not true for the degree that is higher than 3. This paper provides a set of counterexamples of degree 4 [28].

6.1.15. **Registration of point clouds using sample-sphere and adaptive distance restriction**

**Participants:** Yu Meng, Hui Zhang.

Registration of point clouds is a fundamental problem in shape acquisition and shape modeling. In this paper, a novel technique, the sample-sphere method, is proposed to register a pair of point clouds in arbitrary initial positions. This method roughly aligns point clouds by matching pairs of triplets of points, which are approximately congruent under rigid transformation. For a given triplet of points, this method can find all its approximately congruent triplets in $O(kn \log n)$ time, where $n$ is the number of points in the point cloud, and $k$ is a constant depending only on a given tolerance to the rotation error. By employing the techniques of wide bases and largest common point set (LCP), our method is resilient to noise and outliers. Another contribution of this paper is proposing an adaptive distance restriction to improve ICP (iterative closest point) algorithm, which is a classical method to refine rough alignments. With this restriction, the improved ICP is able to reject unreasonable corresponding point pairs during each iteration, so it can precisely align the point clouds which have large non-overlapping regions [24].

6.1.16. **Torus/torus intersection**

**Participants:** Xiao-Ming Liu, Chang-Yuan Liu, Jun-Hai Yong, Jean-Claude Paul.

This paper presents a new algorithm for torus/torus intersection. The pre-image of the intersection in the parametric space of one torus is represented by an implicit equation. The pre-image is divided into one-valued function curve segments by characteristic points. The topological feature of these characteristic points is analyzed to obtain the structure of the pre-image. Intersection curves satisfying required precision are found by a self-adaptive refinement method. Experiment results are presented to illustrate the stability and efficiency of the method [21].

6.1.17. **Polyline approach for approximating Hausdorff distance between planar free-form curves**

**Participants:** Yan-Bing Bai, Jun-Hai Yong, Jean-Claude Paul, Xiao-Ming Liu.
This paper presents a practical polyline approach for approximating Hausdorff distance between planar free-form curves. After the input curves are approximated with polylines using recursively splitting method, the precise Hausdorff distance between polylines is computed as approximation of Hausdorff distance between free-form curves, and the error of approximation is controllable. The computation of Hausdorff distance between polylines is based on an incremental algorithm that computes directed Hausdorff distance from a line segment to a polyline. Furthermore, not every segment on polylines contributes to the final Hausdorff distance. Based on the bound properties of Hausdorff distance and continuity of polylines, two pruning strategies are applied in order to prune useless segments. R-Tree structure is employed as well to accelerate the pruning process. We experimented our algorithm on sets of Bezier, B-Spline and NURBS curves respectively, and there are 95% segments pruned on approximating polylines in average. Two comparisons are also presented: One is with an algorithm computing directed Hausdorff distance on polylines by building Voronoi diagram of segments, the other is with equation solving method for computing Hausdorff distance between free-form curves [1].

6.1.18. Algorithm for orthogonal projection of parametric curves onto B-spline surfaces

Participants: Hai-Chuan Song, Xiao-Ming Liu, Jean-Claude Paul.

This paper proposes an algorithm for calculating the orthogonal projection of parametric curves onto B-spline surfaces. It consists of a second order tracing method with which we construct a polyline to approximate the pre-image curve of the orthogonal projection curve in the parametric domain of the base surface. The final 3D approximate curve is obtained by mapping the approximate polyline onto the base surface. The Hausdorff distance between the exact orthogonal projection curve and the approximate curve is controlled under the user-specified distance tolerance. And the continuity of the approximate curve is $\varepsilon - G^1$, where $\varepsilon$ is the user-specified angle tolerance. Experiments demonstrate that our algorithm is faster than the existing first order algorithms [8].

6.1.19. Converting sectional views to three orthographic views to reconstruct 3D models

Participants: Fengqing Ding, Hui Zhang, Yamei Wen.

Compared with the CSG-based approach, the Brep-based approach has several advantages to construct 3D models from 2D engineering drawings, such as the structure is simpler and the domain of objects that can be handled is wider. However, this approach cannot handle sectional views directly. In this paper, a new method of converting sectional views to three orthographic views is presented. Firstly, the views which have the same projection direction are merged into one view. If the number of views is two, then a new view will be added according to the coordinate relations. Secondly, elements which have been omitted in sectional views are recovered according to the matching information of the existing edges. Finally, the existing Brep-based approach is used to reconstruct the 3D models. The algorithm can handle full sections, broken-out sections, offset sections as well as two orthographic views. The algorithm has been validated by experiments [15].

6.1.20. Computing the Minimum Distance between Two Tori

Participants: Xiao-Ming Liu, Chang-Yuan Liu, Qiang Hu, Jun-Hai Yong.

The minimal distance computing between two tori is the basis of their collision detection and intersection. A method is proposed for discriminating three types of position relationship (i.e., inclusion, disjunction and intersection) between two tori, and for computing their minimal distance. This paper proves that the Hausdorff distance between two circles in three-dimensional space can be obtained by computing their collinear normal points, which can be calculated by solving an equation of degree 8. With classification and comparison of the collinear normal points, the minimum distance and the Hausdorff distance between these two circles are obtained. In addition, this paper proves that the position relationship between two tori relates to not only the minimum distance but also the directed Hausdorff distance between their central circles. And then the minimum distance between two tori is calculated. Numerical results are presented to illustrate the stability and efficiency of the method [20].
6.1.21. Computing the Inner Distances of Volumetric Models for Articulated Shape Description with a Visibility Graph

**Participants:** Yu-Shen Liu, Karthik Ramani, Min Liu.

A new visibility graph based algorithm is presented for computing the inner distances of a 3D shape represented by a volumetric model. The inner distance is defined as the length of the shortest path between landmark points within the shape. The inner distance is robust to articulation and can reflect well the deformation of a shape structure without an explicit decomposition. Our method is based on the visibility graph approach. To check the visibility between pairwise points, we propose a novel, fast, and robust visibility checking algorithm based on a clustering technique which operates directly on the volumetric model without any surface reconstruction procedure, where an octree is used for accelerating the computation. The inner distance can be used as a replacement for other distance measures to build a more accurate description for complex shapes, especially for those with articulated parts [22].

6.1.22. An extended schema and its production rule-based algorithms for assembly data exchange using IGES

**Participants:** Kai-Mo Hu, Bin Wang, Yong Liu, Jing Huang, Jun-Hai Yong.

Assembly data exchange and reuse play an important role in CAD and CAM in shortening the product development cycle. However, current CAD systems cannot transfer mating conditions via neutral file format, and their exported IGES files are heterogeneous. In this paper, a schema for the full data exchange of assemblies is presented based on IGES. We first design algorithms for the pre-and-post processors of parts based on solid model, in which the topologies are explicitly specified and will be referred by mating conditions, and then extend the IGES schema by introducing the Associativity Definition Entity and Associativity Instance Entity defined in IGES standard, so as to represent mating conditions. Finally, a production rule-based method is proposed to analyze and design the data exchange algorithms for assemblies. Within this schema, the heterogeneous representations of assemblies exported from different CAD systems can be processed appropriately, and the mating conditions can be properly exchanged. Experiments on the prototype system verify the robustness, correctness, and flexibility of our schema [49].

6.1.23. Robust shape normalization of 3D articulated, volumetric models

**Participants:** Chao Wang, Yu-Shen Liu, Min Liu, Jun-Hai Yong, Jean-Claude Paul.

3D shape normalization is a common task in various computer graphics and pattern recognition applications. It aims to normalize different objects into a canonical coordinate frame with respect to rigid transformations containing translation, rotation and scaling in order to guarantee a unique representation. However, the conventional normalization approaches do not perform well when dealing with 3D articulated objects. To address this issue, we introduce a new method for normalizing a 3D articulated object in the volumetric form. We use techniques from robust statistics to guide the classical normalization computation. The key idea is to estimate the initial normalization by using implicit shape representation, which produces a novel articulation insensitive weight function to reduce the influence of articulated deformation. We also propose and prove the articulation insensitivity of implicit shape representation. The final solution is found by means of iteratively reweighted least squares. Our method is robust to articulated deformation without any explicit shape decomposition. The experimental results and some applications are presented for demonstrating the effectiveness of our method [27].


**Participants:** Zhi Chen, Bin Wang, Norbert Muller, Hui Zhang, Jun-Hai Yong.
The summation of floating-point numbers is ubiquitous in computer systems, while computation implemented in fixed length floating-point arithmetic may lead to inaccurate result due to rounding error. This paper presents an efficient algorithm which produces a faithful result by combining splitting the mantissa and error-free accumulation. Each summand is split into several parts with limited significant bits, which ensures these parts can be accumulated without rounding error under certain conditions. In the implementation, we discuss how to get exponent of floating-point number quickly, which is key to decide how to split summand. Our method works on computers complying with IEEE 754 standard. The running time of our algorithm is proportional to the size of the input data, according to both analysis and numerical tests [12].

6.1.25. 3DMolNavi: A navigation system for flexible molecular shape retrieval based on histogram and dimensionality reduction

Participants: Yu-Shen Liu, Meng Wang, Jean-Claude Paul.

3DMolNavi is a web-based visualized navigation system developed for intuitively exploring flexible molecular shape retrieval. This system is based on the histogram of Inner Distance Shape Signature (IDSS) for fast retrieving molecules that are similar to a query molecule, and uses dimensionality reduction to navigate the retrieved results in 2D and 3D spaces [23].


Participants: Hui Kong, Hehua Zhang, Xiaoyu Song, Ming Gu, Jiaguang Sun.

Geometric algorithms are widely used in many scientific fields like computer vision, computer graphics. To guarantee the correctness of these algorithms, it’s important to apply formal method to them. In this paper, we propose an approach to proving the correctness of geometric algorithms. The main contribution of the paper is that a set of proof decomposition rules is proposed which can help improve the automation of the proof of geometric algorithms. We choose TLA+2, a structural specification and proof language, as our experiment environment. The case study on a classical convex hull algorithm shows the usability of the method [36].

6.1.27. Multi-resolution mesh fitting by B-spline surfaces for reverse engineering

Participants: Sen Zhang, Zhigang Li, Hui Zhang.

This paper presents a new multi-resolution mesh fitting algorithm, extending the adaptive patch-based fitting scheme where each underlying quadrilateral is recursively subdivided into four sub-patches. In this paper, the $G^1$ continuity constraints, which mainly consist of perpendicular constraints and twist compatibility constraints, are deduced for B-spline patches. In order to construct a unique B-spline patch for each quadrilateral, the mesh vertices are applied in a least-square approximation, and the energy functions associated with a patch are minimized. In contrast to the original algorithm, this paper fits the mesh into B-spline patches instead of Bézier patches with $G^1$ continuity. The B-spline patches make the algorithm have more free control points to be used for optimizing the shape of the quadrilateral patches to achieve higher flexible patch control and less recursive times [42].

6.1.28. An example-driven symbol recognition approach based on key features in engineering drawings

Participants: Tiantian Guo, Hui Zhang, Yamei Wen.

In this paper, we present an example-driven symbol recognition algorithm based on its key features in CAD engineering drawings. When user provides an example of a specific symbol, the input symbol is analyzed and its features are extracted automatically. Based on the relation representation, the constrained tree with key feature priority can be established for this type of symbol. By this means, the symbol library can be built and expanded automatically in order to handle variety engineering drawings. In the next stage of the recognition processes, we first locate the key feature nodes in drawings, and then find other elements around which satisfy the topology structure of constrained tree. If all the elements and constrains in the tree are found, the symbol object will be recognized. Because of the accurate position, unnecessary matching calculations are greatly reduced. Experimental results validate that our approach is effective [33].
6.1.29. **Automatic generation of canonical views for CAD models**  
**Participants:** Kaimo Hu, Bin Wang, Bin Yuan, Junhai Yong.

Selecting the best views for 3D objects is useful for many applications. However, with the existing methods applied in CAD models, the results neither exhibit the 3D structures of the models fairly nor conform to human’s browsing habits. In this paper, we present a robust method to generate the canonical views of CAD models, and the above problem is solved by considering the geometry and visual salient features simultaneously. We first demonstrate that for a CAD model, the three coordinate axes can be approximately determined by the scaled normals of its faces, such that the pose can be robustly normalized. A graph-based algorithm is also designed to accelerate the searching process. Then, a convex hull based method is applied to infer the upright orientation. Finally, four isometric views are selected as candidates, and the one whose depth image owns the most visual features is selected. Experiments on the Engineering Shape Benchmark (ESB) show that the views generated by our method are pleasant, informative and representative. We also apply our method in the calculation of model rectilinearity, and the results demonstrate its high performance [34].

6.2. Geometric Uncertainties

6.2.1. **G^n blending multiple surfaces in polar coordinates**  

This paper proposes a method of $G^n$ blending multiple parametric surfaces in polar coordinates. It models the geometric continuity conditions of parametric surfaces in polar coordinates and presents a mechanism of converting a Cartesian parametric surface into its polar coordinate form. The basic idea is first to re-parameterize the parametric blendees into the form of polar coordinates. Then they are blended simultaneously by a basis function in the complex domain. To extend its compatibility, we also propose a method of converting polar coordinate blending surface into N NURBS patches. One application of this technique is to fill N-sided holes. Examples are presented to show its feasibility and practicability [6].

6.2.2. **Filling n-sided regions with $G^1$ triangular Coons B-spline patches**  

Filling n-sided regions is an essential operation in shape and surface modeling. Positional and tangential continuities are highly required in designing and manufacturing. We propose a method for filling n-sided regions with untrimmed triangular Coons B-spline patches, preserving $G^1$ continuity exactly. The algorithm first computes a central point, a central normal, the central, and the corner derivative vectors. Then the region is split into n triangular areas by connecting the central point to each corner of the boundary. These inner curves and all cross-boundary derivatives are computed fulfilling $G^1$ compatibility conditions. And finally, the triangular patches are generated in the Coons B-spline form, one boundary of which is regressed to the central vertex. Neither positional nor tangential error is introduced by this method. And only one-degree elevation is needed [54].

6.2.3. **G^n Filling orbicular N-sided holes using periodic B-spline surfaces**  
**Participants:** K.-L. Shi, J.-H. Yong, J.-C. Paul.

The orbicular N-sided hole filling problem is usually introduced by filleting an end-point of a part with large radius. The existing methods based on quadrilateral partition or constrained-optimization can rarely generate high-order continuous blending surfaces under these circumstances. This paper first reparameterizes the boundary of the specified orbicular N-sided hole to ensure the compatibility of neighboring cross-boundary derivatives on the connecting points, preserving their $G^n$ continuity. Then we compute the control points of the periodic B-spline surface using the sufficient $G^n$ continuity condition on the pole and the algorithm of extending parametric surfaces. This method generates single blending surface, which can be converted into standard B spline surface by adding knots without introducing errors. It only elevates the degree of the boundary by n. The construction method is simple and efficient, without iteration nor large-scale matrix solving. It achieves $G^n$ continuity under compatible conditions. The blending examples underline its feasibility and practicability [55].
6.2.4. A Thin-plate CAD Mesh Model Splitting Approach Based on Fitting Primitives
Participants: Chun Geng, Hiromasa Suzuki, Dong-Ming Yan, Takasi Michikawa, Yuichi Sato, Masayoshi Hashima, Eiji Ohta.

Extracting structural information from mesh models is crucial for Simulation Driven Design (SDD) in industrial applications. Focusing on thinplate CAD mesh models (the most commonly used parts in electronic products such as PCs, mobile phones and so on), we present an algorithm based on primitive fitting for segmenting thinplate CAD mesh models into parts of three different types, two of which are extruding surfaces and the other is a lateral surface. This method can be used for solid model reconstruction in the SDD process. Our approach involves two steps. First, a completely automatic method for accurate primitive fitting on CAD meshes is proposed based on the hierarchical primitive fitting framework. In the second step, a novel procedure is proposed for splitting thinplate CAD mesh models by detecting parallel extruding surfaces and lateral surfaces. The method presented here has been proved to work smoothly in applications of real product design [46].

6.2.5. A face-based shape matching method for IGES surface model
Participants: Kaimo Hu, Bin Wang, Yi Gao, Qiming Yuan, Junhai Yong.

IGES is a widely used standard for mechanical data exchange. In this paper, we present a new method for the retrieval task of IGES surface model. Based on this method, a novel distinctive face selection strategy is proposed and evaluated. In the training database, each model is treated as a set of disordered faces, and their features are extracted and stored respectively. The Discounted Cumulative Gain (DCG) value of each face is then calculated and stored for later utilization. To retrieve models in the testing database, we first forecast each face’s DCG value by searching its most similar face’s DCG value in training database, and then the top k faces with highest forecasted DCGs are selected as query input. A greedy algorithm is finally applied to get the total similarity. Experimental results show that our algorithm is superior or at least comparable to some of the most powerful methods in finding parts with similar functionality in most cases [48].

6.2.6. Generating B-spline curves based on control-point interpolation
Participants: Jing Liu, Kan-Le Shi, Jun-Hai Yong, He-Jin Gu.

Generating smooth B-spline curves is a fundamental operation of computer aided geometric design. This paper presents a method to calculate unknown control points using specified control points and knots to generate a smooth B-spline curve. It is based on the basis-function-maximum-value parameterization introduced in this paper. This method first parameterizes all control points; then regards given control points as data points to create a fit curve by interpolation; and finally obtains the unknown control points by evaluating the corresponding parameters directly, which ensures the continuity and smoothness of the generated B-spline curve. The examples in the last section illustrate the feasibility of this method [51].

6.2.7. A new functionality-based benchmark for basic CAD Model retrieval
Participants: Kaimo Hu, Bin Wang, Yi Gao, Dong Li, Junhai Yong.

In this paper, we propose a new functionality-based benchmark for CAD model retrieval. Our benchmark contains 1968 frequently-used CAD models which are divided into training set and test set. The models are carefully classified by their functionalities in industry. Eight different shape descriptors are then compared using four famous evaluation measurements. The results show that models having the same functionalities do not necessarily share the same or similar shapes, hence the functionality-based retrieval methods are encouraged, which we believe will be of great help for the improvements of design reusability. Some possible future work for 3D model retrieval in mechanical domain are also proposed based on the observation of our experiments [47].

6.2.8. Shape similarity assessment approach for CAD models based on graph edit distance
Participants: Bin Wang, Dong Li, Kaimo Hu, Hui Zhang.
This paper proposes a new shape similarity assessment approach for CAD models in Boundary Representation (Brep) based on graph edit distance. A suboptimal computational procedure is performed to find the best alignment between local structures sets of attributed graphs derived from models. Assuming that only a minority of local structures characterize the functionality, we figure out the weight of every local structure in the query model through a training phase, and then evaluate the similarity between two models by calculating the weighted graph edit distance of corresponding attributed graphs. Experiment results show that our method provides solid retrieval performance on a real-world CAD model database [57].

6.2.9. The transition between sharp and rounded features and the manipulation of incompatible boundary in filling n-sided holes

Participants: Kan-Le Shi, Jun-Hai Yong, Peng Liu, Jia-Guang Sun, Jean-Claude Paul.

N-sided hole filling plays an important role in vertex blending. Piegl and Tiller presented an algorithm to interpolate the given boundary and cross-boundary derivatives in B-spline form. To deal with the incompatible cases that their algorithm cannot handle, we propose an extension method to manipulate the transition between sharp and rounded features. The algorithm first patches n crescent-shaped extended surfaces to the boundary with \( G^2 \) continuity to handle incompatibility problem in the corners. Then, we compute the inner curves and the corresponding cross-boundary derivatives fulfilling tangent and twist compatibilities. The generated B-spline Coons patches are \( G^1 \)-continuously connected exactly, and have \( \varepsilon - G^1 \) continuity with the extended surfaces. Our method improves the continuity-quality of the shape and reduces the count of the inserted knots. It can be applied to all \( G^0 \)-continuous boundary conditions without any restrictions imposed on the boundary or cross-boundary derivatives. It generates better shapes than some popular industrial modeling systems on these incompatible occasions. Some examples underline its feasibility [53].

6.2.10. Epsilon-G2 B-spline surface interpolation

Participants: Kan-Le Shi, Jun-Hai Yong, Jia-Guang Sun, Jean-Claude Paul.

This paper proposes a method to construct a B-spline surface that interpolates the specified four groups of boundary derivative curves in the B-spline form. The discontinuity can be bounded by an arbitrary geometric invariant as the tolerance. The method first handles the six types of the compatibility problems by continuity-preserving reparameterization, knot-insertion and local control-point tuning. The transformed boundary conditions are then parametrically compatible, so the Coons strategy can be applied to construct the final interpolant. Not only can it be used in the reliable geometric modeling, but the approach also can be applied to many other algorithms that require compatibility guarantee [26].

6.2.11. \( G^2 \) B-spline interpolation to a closed mesh


This paper focuses on interpolating vertices and normal vectors of a closed quad-dominant mesh \( 1G^2 \)-continuously using regular Coons B-spline surfaces, which are popular in industrial CAD/CAM systems. We first decompose all non-quadrangular facets into quadrilaterals. The tangential and second-order derivative vectors are then estimated on each vertex of the quads. A least-square adjustment algorithm based on the homogeneous form of \( G^2 \) continuity condition is applied to achieve curvature continuity. Afterwards, the boundary curves, the first- and the second-order cross-boundary derivative curves are constructed fulfilling \( G^2 \) continuity and compatibility conditions. Coons B-spline patches are finally generated using these curves as boundary conditions. In this paper, the upper bound of the rank of \( G^2 \) continuity condition matrices is also strictly proved to be 2n-3, and the method of tangent-vector estimation is improved to avoid petal-shaped patches in interpolating solids of revolution [7].

6.2.12. A new method for identifying and validating features from 2D sectional views

Participants: Yamei Wen, Hui Zhang, Jiaguang Sun, Jean-Claude Paul.
Feature identification is one of the key steps for 3D solids reconstruction from 2D vector engineering drawings using the volume-based method. In this paper, we propose a novel method to identify and validate features from sectional views. First, features are classified as explicit features (EPFs) and implicit features (IPFs), which are then identified in an order of priority using heuristic hints. We show that the problem of constructing EPFs can be formulated as a 0-1 integer linear program (ILP), and the IPFs are generated based on the understanding of semantic information of omitted projections in sectional views. Then, the Loop-Relation Graph (LRG) is introduced as a multi-connected-subgraph representation for describing the relations between loops and features. According to the LRG, a reasoning technique based on confidence is implemented to interactively validate features. This method can recover features without complete projections, and the level of understanding sectional views is improved. Full sections, partial sections, offset sections as well as revolved sections can be handled by our method. Several examples are provided to demonstrate the practicability of our approach [29].

6.2.13. Ridge extraction of a smooth 2-manifold surface based on vector field
Participants: Wujun Che, Xiaopeng Zhang, Yi-Kuan Zhang, Jean-Claude Paul.

This paper presents a general scheme to compute ridges on a smooth 2-manifold surface from the standpoint of a vector field. A ridge field is introduced. Starting with an initial ridge, which may or may not be umbilical, a ridge line is then traced by calculating an associated integral curve of this field in conjunction with a new projection procedure to prevent it from diverging. This projection is the first that can optimize a ridge guess to lie on a ridge line uniquely and accurately. In order to follow this scheme, we not only develop practical ridge formulae but also address their corresponding computational procedures for an analytical surface patch, especially for an implicit surface. In contrast to other existing methods, our new approach is mathematically sound and characterized by considering the full geometric structures and topological patterns of ridges on a generic smooth surface. The resulting ridges are accurate in the numerical sense and meet the requirement of high accuracy with complete topology. Although the objective of this paper is to develop a mathematically sound framework for ridges on a smooth surface, we give a comprehensive review of relevant works on both meshes and smooth surfaces for readers [11].

6.2.14. Manifold-ranking based retrieval using $k$-regular nearest neighbor graph
Participants: Bin Wang, Feng Pan, Kaimo Hu, Jean-Claude Paul.

Manifold-ranking is a powerful method in semi-supervised learning, and its performance heavily depends on the quality of the constructed graph. In this paper, we propose a novel graph structure named $k$-regular nearest neighbor ($k$-RNN) graph as well as its constructing algorithm, and apply the new graph structure in the framework of manifold-ranking based retrieval. We show that the manifold-ranking algorithm based on our proposed graph structure performs better than that of the existing graph structures such as $k$-nearest neighbor ($k$-NN) graph and connected graph in image retrieval, 2D data clustering as well as 3D model retrieval. In addition, the automatic sample reweighting and graph updating algorithms are presented for the relevance feedback of our algorithm. Experiments demonstrate that the proposed algorithm outperforms the state-of-the-art algorithms [9].

6.2.15. Efficient computation of clipped Voronoi diagram for mesh generation
Participants: Dong-Ming Yan, Wenping Wang, Bruno Lévy, Yang Liu.

The Voronoi diagram is a fundamental geometric structure widely used in various fields, especially in computer graphics and geometry computing. For a set of points in a compact domain (i.e. a bounded and closed 2D region or 3D volume), some Voronoi cells of their Voronoi diagram are infinite or partially outside of the domain, but in practice only the parts of the cells inside the domain are needed, as when computing the centroidal Voronoi tessellation. Such a Voronoi diagram confined to a compact domain is called a clipped Voronoi diagram. We present an efficient algorithm for computing the clipped Voronoi diagram for a set of sites with respect to a compact 2D region or a 3D volume. We also apply the proposed method to optimal mesh generation based on the centroidal Voronoi tessellation [30].
6.2.16. Automatic Generation of Canonical Views for CAD Models

Participants: Kaimo Hu, Bin Wang, Bin Yuan, Junhai Yong.

Selecting the best views for 3D objects is useful for many applications. However, with the existing methods applied in CAD models, the results neither exhibit the 3D structures of the models fairly nor conform to human’s browsing habits. In this paper, we present a robust method to generate the canonical views of CAD models, and the above problem is solved by considering the geometry and visual salient features simultaneously. We first demonstrate that for a CAD model, the three coordinate axes can be approximately determined by the scaled normals of its faces, such that the pose can be robustly normalized. A graph-based algorithm is also designed to accelerate the searching process. Then, a convex hull based method is applied to infer the upright orientation. Finally, four isometric views are selected as candidates, and the one whose depth image owns the most visual features is selected. Experiments on the Engineering Shape Benchmark (ESB) show that the views generated by our method are pleasant, informative and representative. We also apply our method in the calculation of model rectilinearity, and the results demonstrate its high performance [35].

6.2.17. Mechanical Parts Retrieval Based on Typical Face Matching

Participants: Yi Gao, Bin Wang, Kaimo Hu, Junhai Yong.

This paper presents a face-based retrieval algorithm to search mechanical parts with similar partial features. The method makes it easier to retrieve models with partial features so as to support early stage reusability. In the training phase, all the faces in the database are trained and assigned with a value indicating their distinction. Trivial faces and atypical ones are removed in this phase to improve online retrieval efficiency. In the query phase, we evaluate the distinction of the input faces by aligning them with faces in the database. A greedy algorithm is finally applied to match the input faces and the faces in the database according to their similarity order. Experimental results show that our method can provide a favorable performance when applied to retrieve the models with common partial features comparing to some other mesh-based methods [18].

6.2.18. Continuity Transition with a Single Regular Curved-Knot B-Spline Surface

Participants: Kan-Le Shi, Jun-Hai Yong, Jean-Claude Paul, Jia-Guang Sun.

We propose a canonical form of the curved-knot B-spline surface called the regular curved-knot B-spline. On one hand it allows the transition of the knot vectors so that the continuity configurations of the two opposite boundaries can be different. On the other hand, the regular form achieves the simplicity in storage, evaluation and the construction algorithms, and that makes it possible to be applied in the industrial geometric modeling systems. The applications: bridging, multi-sided hole filling and irregular feature modeling, show that it is well suited for modeling complicated objects, such as a transition between sharp and rounded features. Compared with patching numbers of B-splines, it not only increases the inter-surface continuity of the shape, but also reduces the complexity of algorithms [25].

6.2.19. Meshless quadrangulation by global parameterization

Participants: Er Li, Bruno Lévy, Xiaopeng Zhang, Wu-Jun Che, Weiming Dong, Jean-Claude Paul.

Point cloud is a basic description of discrete shape information. Parameterization of unorganized points is important for shape analysis and shape reconstruction of natural objects. In this paper we present a new algorithm for global parameterization of an unorganized point cloud and its application to the meshing of the cloud. Our method is guided by principal directions so as to preserve the intrinsic geometric properties. After initial estimation of principal directions, we develop a kNN(k-nearest neighbor) graph-based method to get a smooth direction field. Then the point cloud is cut to be topologically equivalent to a disk. The global parameterization is computed and its gradients align well with the guided direction field. A mixed integer solver is used to guarantee a seamless parameterization across the cut lines. The resultant parameterization can be used to triangulate and quadrangulate the point cloud simultaneously in a fully automatic manner, where the shape of the data is of any genus [19].
6.3. Computer Graphics

6.3.1. Multi-Image Based Photon Tracing for Interactive Global Illumination of Dynamic Scenes

Participants: Chunhui Yao, Bin Wang, Bin Chan, Junhai Yong, Jean-Claude Paul.

Image space photon mapping has the advantage of simple implementation on GPU without pre-computation of complex acceleration structures. However, existing approaches use only a single image for tracing caustic photons, so they are limited to computing only a part of the global illumination effects for very simple scenes. In this paper we fully extend the image space approach by using multiple environment maps for photon mapping computation to achieve interactive global illumination of dynamic complex scenes. The two key problems due to the introduction of multiple images are 1) selecting the images to ensure adequate scene coverage; and 2) reliably computing ray-geometry intersections with multiple images. We present effective solutions to these problems and show that, with multiple environment maps, the image-space photon mapping approach can achieve interactive global illumination of dynamic complex scenes. The advantages of the method are demonstrated by comparison with other existing interactive global illumination methods [10].

6.3.2. Quality Solid Texture Synthesis using Position and Index Histogram Matching

Participants: Jiating Chen, Bin Wang.

The synthesis quality is one of the most important aspects in solid texture synthesis algorithms. In recent years several methods are proposed to generate high quality solid textures. However, these existing methods often suffer from the synthesis artifacts such as blurring, missing texture structures, introducing aberrant voxel colors, and so on. In this paper, we introduce a novel algorithm for synthesizing high quality solid textures from 2D exemplars. We first analyze the relevant factors for further improvements of the synthesis quality, and then adopt an optimization framework with the k-coherence search and the discrete solver for solid texture synthesis. The texture optimization approach is integrated with two new kinds of histogram matching methods, position and index histogram matching, which effectively cause the global statistics of the synthesized solid textures to match those of the exemplars. Experimental results show that our algorithm outperforms or at least is comparable to the previous solid texture synthesis algorithms in terms of the synthesis quality [44].

6.3.3. Real-time rendering of heterogeneous translucent objects with arbitrary shapes

Participants: Yajun Wang, Jiaping Wang, Nicolas Holzschuch, Kartic Subr, Jun-Hai Yong, Baining Guo.

We present a real-time algorithm for rendering translucent objects of arbitrary shapes. We approximate the scattering of light inside the objects using the diffusion equation, which we solve on-the-fly using the GPU. Our algorithm is general enough to handle arbitrary geometry, heterogeneous materials, deformable objects and modifications of lighting, all in real-time. In a pre-processing step, we discretize the object into a regular 4-connected structure (QuadGraph). Due to its regular connectivity, this structure is easily packed into a texture and stored on the GPU. At runtime, we use the QuadGraph stored on the GPU to solve the diffusion equation, in real-time, taking into account the varying input conditions: Incoming light, object material and geometry. We handle deformable objects, provided the deformation does not change the topological structure of the objects [58].

6.3.4. Fast Local Color Transfer via Dominant Colors Mapping

Participants: Weiming Dong, Guanbo Bao, Xiaopeng Zhang, Jean-Claude Paul.

We present a novel algorithm to address the above issues. Our method establishes a tight connection between the local color statistics of the source and target images. All the obvious color features can be presented in the result [45].

6.3.5. Real-time watercolor illustrations and animation on GPU

This paper presents a real-time approach to render 3D scenes with the effects of watercolor on GPU. Most processes of the approach are implemented with image-space techniques. Our algorithm renders detail layer, ambient layer and stroke layer separately, and then combines them into final result. During the rendering processes, we use screen space ambient occlusion and shadow mapping to compute shadow in much shorter time, and we use image filter approach to simulate important effects of watercolor. Because our approach is mainly implemented with image-space techniques, it is convenient to use GPU to accelerate the rendering processes and finally our approach achieves real-time speed [59].

6.3.6. A Hierarchical Grid Based Framework for Fast Collision Detection


We present a novel hierarchical grid based method for fast collision detection (CD) for deformable models on GPU architecture. A two-level grid is employed to accommodate the non-uniform distribution of practical scene geometry. A bottom-to-top method is implemented to assign the triangles into the hierarchical grid without any iteration while a deferred scheme is introduced to efficiently update the data structure. To address the issue of load balancing, which greatly influences the performance in SIMD parallelism, a propagation scheme which utilizes a parallel scan and a segmented scan is presented, distributing workloads evenly across all concurrent threads. The proposed method supports both discrete collision detection (DCD) and continuous collision detection (CCD) with self-collision. Some typical benchmarks are tested to verify the effectiveness of our method. The results highlight our speedups over prior algorithms on different commodity GPUs [17].

6.3.7. Improved Stochastic Progressive Photon Mapping with Metropolis Sampling

Participants: Jiating Chen, Bin Wang, Junhai Yong.

This paper presents an improvement to the stochastic progressive photon mapping (SPPM), a method for robustly simulating complex global illumination with distributed ray tracing effects. Normally, similar to photon mapping and other particle tracing algorithms, SPPM would become inefficient when the photons are poorly distributed. An inordinate amount of photons are required to reduce the error caused by noise and bias to acceptable levels. In order to optimize the distribution of photons, we propose an extension of SPPM with a Metropolis-Hastings algorithm, effectively exploiting local coherence among the light paths that contribute to the rendered image. A well-designed scalar contribution function is introduced as our Metropolis sampling strategy, targeting at specific parts of image areas with large error to improve the efficiency of the radiance estimator. Experimental results demonstrate that the new Metropolis sampling based approach maintains the robustness of the standard SPPM method, while significantly improving the rendering efficiency for a wide range of scenes with complex lighting [14].

6.3.8. Efficient Depth-of-Field Rendering with Adaptive Sampling and Multiscale Reconstruction


Depth-of-field is one of the most crucial rendering effects for synthesizing photorealistic images. Unfortunately, this effect is also extremely costly. It can take hundreds to thousands of samples to achieve noise-free results using Monte Carlo integration. This paper introduces an efficient adaptive depth-of-field rendering algorithm that achieves noise-free results using significantly fewer samples. Our algorithm consists of two main phases: adaptive sampling and image reconstruction. In the adaptive sampling phase, the adaptive sample density is determined by a ‘blur-size’ map and ‘pixel-variance’ map computed in the initialization. In the image reconstruction phase, based on the blur-size map, we use a novel multiscale reconstruction filter to dramatically reduce the noise in the defocused areas where the sampled radiance has high variance. Because of the efficiency of this new filter, only a few samples are required. With the combination of the adaptive sampler and the multiscale filter, our algorithm renders near-reference quality depth-of-field images with significantly fewer samples than previous techniques [13].

6.3.9. Real-Time Illumination of Complex Lights Using Divided Sampling

Participants: Chunhui Yao, Bin Wang, Junhai Yong.
Existing methods for real-time illumination of complex lights, either require long time pre-computation, or only focus on some special types of illumination. Because computing different kinds of illumination requires different sampling strategies, this paper introduces a novel efficient framework for rendering illumination of complex light sources, in which the pre-computation is not necessary. We divide the rendering equation into three parts: high-frequency term, low-frequency term and occlusion term. Each term is computed by a proper sampling strategy. High-frequency term is solved by importance sampling the BRDF, while low-frequency term is computed by importance sampling the light sources. Occlusion term is computed with depth information in screen-space, and the required number of samples is greatly reduced by interleaved sampling. Our framework is easy to implement on GPU and can solve many real-time rendering problems. We take real-time environment-map-lighting as an example for demonstrating applications of this framework. The results show that our technique can handle complete light effects with higher quality than previous works [31].

6.3.10. Fast Multi-Operator Image Resizing and Evaluation
Participants: Weiming Dong, Guanbo Bao, Xiaopeng Zhang,Jean-Claude Paul.

Current multi-operator image resizing methods succeed in generating impressive results by using image similarity measure to guide the resizing process. An optimal operation path is found in the resizing space. However, their slow resizing speed caused by inefficient computation strategy of the bidirectional patch matching becomes a drawback for practical use. In this paper, we present a novel method to address this problem. By combining seam carving with scaling and cropping, our method can realize content-aware image resizing very fast. We define cost functions combing image energy and dominant color descriptor for all the operators to evaluate the damage to both local image content and global visual effect. Therefore our algorithm can automatically find an optimal sequence of operations to resize the image by dynamic programming or greedy algorithm. We also extend our algorithm to indirect image resizing which can protect the aspect ratio of the dominant object in an image [16].

6.3.11. Real-time Volume Caustics with Image-based Photon Tracing
Participants: Yuxiang Wang, Bin Wang, Li Chen.

Rendering of volume caustics in participating media is often expensive, even with different acceleration approaches. Basic volume photon tracing is used to render such effects, but rather slow due to its massive quantity of photons to be traced. In this paper we present an image-based volume photon tracing method for rendering volume caustics at real-time frame rates. Motivated by multi-image based photon tracing, our technique uses multiple depth maps to accelerate the intersection test procedure, achieving a plausible and fast rendering of volume caustics. Each photon dynamically selects the depth map layer for intersection test, and the test converges to an approximate solution using image space methods in a few recursions. This allows us to compute photon distribution in participating media while avoiding massive computation on accurate intersection tests with scene geometry. We demonstrate that our technique, combined with photon splatting techniques, is able to render volume caustics caused by multiple refractions [39].

6.3.12. Parallel Spatial Hashing for Collision Detection of Deformable Surfaces

We present a fast collision detection method for deformable surfaces with parallel spatial hashing on GPU architecture. The efficient update and access of the uniform grid are exploited to accelerate the performance in our method. To deal with the inflexible memory system, which makes the building of stream data a challenging task on GPU, we propose to subdivide the whole workload into irregular segments and design an efficient evaluation algorithm, which employs parallel scan and stream compaction, to build the stream data in parallel. The load balancing is a key aspect that needs to be considered in the SIMD parallelism. We break the heavy and irregular collision computation down into lightweight part and heavyweight part, ensuring the later perfectly run in load balancing manner with each concurrent thread processes just a single collision. In practice, our approach can perform collision detection in tens of milliseconds on a PC with NVIDIA GTX 260 graphics card on benchmarks composed of millions of triangles. The results highlight our speedups over prior CPU-based and GPU-based algorithms [32].
6.3.13. Distribution-Aware Image Color Transfer  
**Participants:** Fuzhang Wu, Weiming Dong, Xing Mei, Xiaopeng Zhang, Xiaohong Jia, Jean-Claude Paul.

Color transfer is a practical image editing technology which is useful in various applications. An ideal color transfer algorithm should keep the scene in the source image and apply the color styles of the reference image. All the dominant color styles of the reference image should be presented in the result especially when there are similar contents in the source and reference images. We propose a robust color transfer framework to address the above issues. Our method can establish a soft connections between the local color statistics of the source and reference images. All the obvious color features can be presented in the result image, as well as the spatial distribution of the reference color pattern [40].

6.3.14. Translucent Material Transfer Based on Single Images  
**Participants:** Chao Li, Weiming Dong, Ning Zhou, Xiaopeng Zhang, Jean-Claude Paul.

Extraction and re-rendering of real materials give large contributions to various image-based applications. As one of the key properties of modeling the appearance of an object, materials mainly focus on the effects caused by light transportation. Therefore, understanding the characteristics of a complex material from a single photograph and transferring it to an object in another image becomes a very challenging problem. In this paper, we present a novel framework to transfer real translucent materials such as fruits and flowers between single images. We define a group of information which can model the attributes during the extraction and transfer process. Once we extract this information from both the source and target images, we can easily produce a realistic photograph of an object with target-like materials and suitable shading effects in the environment of sources [37].

6.3.15. Multicage Image Deformation On GPU  
**Participants:** Weiliang Meng, Xiaopeng Zhang, Weiming Dong, Jean-Claude Paul.

As a linear blending method, cage-based deformation is widely used in various applications of image and geometry processing. In most cases especially in the interactive mode, deformation based on embedded cages does not work well as some of the coefficients are not continual and make the deformation discontinuous, which means existing “spring up” phenomenon. However, it’s common for us to deform the ROI (Region of Interest) while keeping local part untouched or with only small adjustments. In this paper, we design a scheme to solve the above problem. A multicage can be generated manually or automatically, and the image deformation can be adjusted intelligently according to the local cage shape to preserve important details. On the other hand, we don’t need to care about the pixels’ position relative to the multicage. All the pixels go through the same process, and this will save a lot of time. We also design a packing method for cage coordinates to pack all the necessary coefficients into one texture. A vertex shader can be used to accelerate the deformation process, leading to realtime deformation even for large images [38].

6.3.16. Real-time volume caustics with image-based photon tracing  
**Participants:** Yuxiang Wang, Bin Wang, Li Chen.

Rendering of volume caustics in participating media is often expensive, even with different acceleration approaches. Basic volume photon tracing is used to render such effects, but rather slow due to its massive quantity of photons to be traced. In this paper we present an image-based volume photon tracing method for rendering volume caustics at real-time frame rates. Motivated by multi-image based photon tracing, our technique uses multiple depth maps to accelerate the intersection test procedure, achieving a plausible and fast rendering of volume caustics. Each photon dynamically selects the depth map layer for intersection test, and the test converges to an approximate solution using image space methods in a few recursions. This allows us to compute photon distribution in participating media while avoiding massive computation on accurate intersection tests with scene geometry. We demonstrate that our technique, combined with photon splatting techniques, is able to render volume caustics caused by multiple refractions [39].

6.3.17. Interactive Visual Simulation of Dynamic Ink Diffusion Effects  
**Participants:** Shibiao Xu, Xing Mei, Weiming Dong, Zhiyi Zhang, Xiaopeng Zhang.
This paper presents an effective method that simulates the ink diffusion process with visual plausible effects and real-time performance. Our algorithm updates the dynamic ink volume with a hybrid grid-particle representation: the fluid velocity field is calculated with a low-resolution grid structure, while the highly detailed ink effects are controlled and visualized with the particles. We propose an improved ink rendering method with particle sprites and motion blur techniques. The simulation and the rendering processes are efficiently implemented on graphics hardware for interactive frame rates. Compared to traditional simulation methods that treat water and ink as two mixable fluids, our method is simple but effective: it captures various ink effects such as pinned boundary [Chu and Tai 2005] and filament pattern [Shiny et al. 2010] with real-time performance; it allows easy interaction with the artists; it includes basic solid-fluid interaction. We believe that our method is attractive for industrial animation and art design [ 41 ].
5. New Results

5.1. Contributions earlier to 2011 but only published in 2011

Participants: Gérard Biau, Olivier Catoni, Sébastien Gerchinovitz, Vincent Rivoirard, Gilles Stoltz.

We do not discuss here the contributions provided by [12], [13], [14], [17], [18], [20], [22], [23], [25], since they were achieved in 2009 or earlier (but only published this year due to long queues in publication tracks of journals). [32] was revised but is still under review.

5.2. Sparse regression estimation

Participants: Gérard Biau, Olivier Catoni, Sébastien Gerchinovitz, Vincent Rivoirard, Gilles Stoltz, Jia Yuan Yu.

Sébastien Gerchinovitz and Jia Yuan Yu continued the work initiated by the former in the above-mentioned conference paper [25]; they derived from the sparsity results in individual sequences presented therein the minimax optimal rates of aggregation for individual sequences on $\ell^1$ balls. In particular, they exhibited, in certain cases, a phase transition between the $\ln T$ and the $\sqrt{T}$ behavior of the minimax regret, where $T$ denotes the number of instances. These results and all previous ones are summarized in the PhD thesis [10].

Other results were obtained in a stochastic framework, where input–output pairs are given by i.i.d. variables; they are described in the technical report [30]. Let $(X,Y)$ be a random pair taking values in $\mathbb{R}^p \times \mathbb{R}$. In the so-called single-index model, one has $Y = f^*(\theta^T X) + W$, where $f^*$ is an unknown univariate measurable function, $\theta^*$ is an unknown vector in $\mathbb{R}^d$, and $W$ denotes a random noise satisfying $\mathbb{E}[W|X] = 0$. The single-index model is known to offer a flexible way to model a variety of high-dimensional real-world phenomena. However, despite its relative simplicity, this dimension reduction scheme faces severe complications as soon as the underlying dimension becomes larger than the number of observations and this is why this estimation problem was considered from a sparsity perspective using a PAC-Bayesian approach.

Last but not least, we mention the edited book [29], which provides a modern overview on high-dimensional estimation.

5.3. Sequential learning with limited feedback; in particular, bandit problems

Participants: Gilles Stoltz, Jia Yuan Yu.

Some of the results cited below are summarized or stated as open problems in the habilitation thesis [11].

5.3.1. Bandit problems

We achieved three contributions. The first is described in the conference paper [27]: it revisits asymptotically optimal results of Lai and Robbins, Burnetas and Katehakis in a non-asymptotic way. The second is stated in the journal article [19] and is concerned with obtaining fast convergence rates for the regret in case of a continuum of arms (of course under some regularity and topological assumptions on the mean-payoff function $f$).

The third one is detailed in [24] and started from the following observation. Typical results in the bandit literature were of the following form: if the regularity of the mean-payoff function $f$ is known (or if a bound on it is known) then the regret is small. Actually, results were usually taking the following weaker form: when the algorithm is tuned with some parameters, then the regret is small against a certain class of stochastic environments. The question was thus to have an adaptive procedure, that, given one unknown environment (with unknown regularity), ensures that the regret is asymptotically small; even better, the desired aim was to control the regret in some uniform manner (in a distribution-free sense up to the regularity parameters). As described in this conference paper, a solution was achieved in the case of Lipschitz environments.
5.3.2. Approachability in games with partial monitoring

The conference paper [28] explains how we could re-obtain, in a simple, more straightforward, and computationally efficient manner a result proven by Perchet in his PhD thesis: the necessary and sufficient condition for the approachability of a closed convex set under partial monitoring.

5.4. Inference

Participant: Gérard Biau.

5.4.1. Geometric inference

This line of research is in collaboration with the Geometrica project-team (INRIA Saclay). As the latter says:

Due to the fast evolution of data acquisition devices and computational power, scientists in many areas are demanding efficient algorithmic tools for analyzing, manipulating and visualizing more and more complex shapes or complex systems from approximating data. Many of the existing algorithmic solutions which come with little theoretical guarantees provide unsatisfactory and/or unpredictable results. Since these algorithms take as input discrete geometric data, it is mandatory to develop concepts that are rich enough to robustly and correctly approximate continuous shapes and their geometric properties by discrete models. Ensuring the correctness of geometric estimations and approximations on discrete data is a sensitive problem in many applications.

Thus, motivated by a broad range of potential applications in topological and geometric inference, we introduce in [15] a weighted version of the $k$-nearest neighbor density estimator. Various pointwise consistency results of this estimator are established; the proposed method is also implemented to recover level sets in both simulated and real-life data.

Another problem of geometric inference is the following one, studied in [16]. Principal curves are nonlinear generalizations of the notion of first principal component. Roughly, a principal curve is a parameterized curve in $\mathbb{R}^d$ that passes through the “middle” of a data cloud drawn from some unknown probability distribution. Depending on the definition, a principal curve relies on some unknown parameters (number of segments, length, turn...) which have to be properly chosen to recover the shape of the data without interpolating. In this paper, we consider the principal curve problem from an empirical risk minimization perspective and address the parameter selection issue using the point of view of model selection via penalization. We offer oracle inequalities and implement the proposed approaches to recover the hidden structures in both simulated and real-life data.

5.4.2. Statistical inference

We still keep an eye on more traditional mathematical statistics; in particular, the technical report [31] takes place within this field. It shows, for a large class of distributions and large samples, that estimates of the variance $\sigma^2$ and of the standard deviation $\sigma$ are more often Pitman closer to their target than the corresponding shrinkage estimates which improve the mean squared error. The results thus indicate that Pitman closeness criterion, despite its controversial nature, should be regarded as a useful and complementary tool for the evaluation of estimates of $\sigma^2$ and of $\sigma$.

5.5. Statistical inference for biological systems based on a size-structured population

Participant: Vincent Rivoirard.
The journal paper [21] considers the problem of estimating the division rate of a size-structured population in a nonparametric setting. The size of the system evolves according to a transport-fragmentation equation: each individual grows with a given transport rate, and splits into two offsprings of the same size, following a binary fragmentation process with unknown division rate that depends on its size. In contrast to a deterministic inverse problem approach, this paper takes the perspective of statistical inference: the data consists in a large sample of the size of individuals when the evolution of the system is close to its time-asymptotic behavior, so that it can be related to the eigenproblem of the considered transport-fragmentation equation. By estimating statistically each term of the eigenvalue problem and suitably inverting a certain linear operator, it constructs a more realistic estimator of the division rate that achieves the same optimal error bound as in related deterministic inverse problems. The procedure relies on kernel methods with automatic bandwidth selection. It is inspired by model selection and recent results of Goldenschluger and Lepski.
3. New Results

3.1. 3D matrix-free $P^1$-exact conservative interpolation

Participants: F. Alauzet [correspondent]

The interpolation stage is a crucial step for time-dependent mesh adaptive simulations. Indeed, if the accuracy of the solution is spoiled during this stage then it is lost for ever. In the past, we have demonstrated the superiority of conservative interpolation for 2D compressible flow simulations. Here, we develop the 3D extension of the method on tetrahedral meshes. The same conclusion arose. The main difficulty was to design a fast, accurate and robust mesh intersection algorithm.

3.2. A changing-topology ALE numerical scheme

Participants: F. Alauzet [correspondent] and G. Olivier

The main difficulty arising in numerical simulations with moving geometries is to handle the displacement of the domain boundaries, i.e., the moving bodies. Only vertices displacement is not sufficient to achieve complex movement such as shear. We proved that the use of edge swapping allows us to achieve such complex displacement. We therefore developed an ALE formulation of this topological mesh modification to preserve the solver accuracy and convergence order [31].

3.3. Advanced boundary layer meshing

Participants: F. Alauzet [correspondent], J. Castelneau, L. Marechal, D. Marcum and A. Loseille

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.

3.4. Applications du maillage à la cryptographie

Participants: T. Grosges [correspondent], D. Barchiesi, Michael François


3.5. Applications du maillage à l’électromagnétisme et modélisation multi-physiques

Participants: D. Barchiesi [correspondant], T. Grosges, Houman Bourouchaki, Laurence Giraud-Moreau, Sameh Kessentini, Anis Chaari

3.6. Automatic decomposition of discretized surfaces for parallel processing

**Participants**: S.H. Lo, H. Borouchaki [correspondant], P. Laug

Free-form surfaces and industrial surface forms could nowadays be conveniently generated efficiently by laser-based digitizing techniques or from a CAD graphics system. Automatic algorithms are imperative to decompose complex triangulated objects consisting of thousands to millions of nodal points into simpler surface parts for parameterization and parallel processing. Non-manifold complicated discretized objects will be handled and simple closed or open surface parts (manifolds) are retrieved by means of pure topological considerations. Each of the simple surface part, which is topologically equivalent to an open surface, a sphere or a torus will be decomposed by entirely topological operations into n equal pieces based on a specified geometrical criterion such as surface area, Gaussian curvatures or magnitudes of dihedral angles, etc. Cut planes could be conveniently defined normal to the axes of inertia of the object to be decomposed. The cut plane which produces surface parts which best respect the given criterion will be chosen, and the two bisected surfaces will be made as equal as possible by means of some general balancing mechanisms. This procedure could be repeated as often as necessary until sub-surface parts satisfying the given criterion are obtained. Each piece of the decomposed surfaces having an intact topological boundary can then be processed independently in parallel, and all the pieces could be put back together to recover the original object or partially to represent a certain portion of the object [29].

3.7. Construction de maillages de degré 2 – Triangle et tétraèdre P2

**Participants**: P.L. George [correspondant], H. Borouchaki, P. Laug

There is a need for finite elements of degree 2 or more to solve various P.D.E. problems. This study discusses a method to construct such meshes in the case of triangular element (in the plane or for a surface) or tetrahedral element (in the volume case), restricting at degree 2. This first part considers the planar case and, to begin with, returns to Bézier curves and Bézier triangles of degree 2. In the case of triangles, the relation with Lagrange P2 finite element is shown. Validity conditions are discussed and some invalid elements are shown while proposing a method to correct them. A construction method is then proposed [34].

3.8. Continuous and Discrete Adjoints to the Euler Equations for Fluids

**Participants**: F. Alauzet [correspondant] and O. Pironneau

Adjoints are used in optimization to speed-up computations, simplify optimality conditions or compute sensitivities. Because time is reversed in adjoint equations with first order time derivatives, boundary conditions and transmission conditions through shocks can be difficult to understand. In this work, we analyzed the adjoint equations that arise in the context of compressible flows governed by the Euler equations of fluid dynamics. We showed that the continuous adjoints and the discrete adjoints computed by automatic differentiation agree numerically; in particular the adjoint is found to be continuous at the shocks and usually discontinuous at contact discontinuities by both [9].

3.9. Goal-oriented anisotropic mesh adaptation for unsteady problems

**Participants**: F. Alauzet [correspondant], A. Belme and A. Dervieux

We have extended our previous work on goal-oriented mesh adaptation to time dependent simulations. This requires to set up a global fixed point algorithm in which state and adjoint variables problems are solved. The adjoint problem is solved backward in time [24].

3.10. Hessian recovery techniques on isotropic and anisotropic meshes

**Participants**: M. Picasso, F. Alauzet [correspondant], H. Borouchaki and P.-L. George
Mesh adaptation required the use of second order derivatives, i.e., the Hessian. For second order numerical scheme, the provided numerical solution is only piecewise linear. Consequently, numerical methods are considered to recover second order derivatives, the famous recovery techniques. In this work, we have investigated several methods. Numerical results on 2D and 3D isotropic and anisotropic meshes indicate that the quality of the results is strongly linked to the mesh topology and that no convergence can be insured in general [22].

3.11. High Quality Geometric Meshing of CAD Surfaces

Participants: P. Laug [correspondant], H. Borouchaki

We propose a general scheme of an indirect approach for generating isotropic and anisotropic geometric meshes of a surface constituted by a conformal assembly of parametric patches, based on the concept of metric. The different steps of the scheme are considered and, in particular, the definition of the geometric metric at each point of the surface (internal to a patch, belonging to an interface or boundary curve, or extremity of such a curve) as well as its corresponding induced metric in parametric domains.

Isotropic or anisotropic geometric metrics can locally produce significant size variations (internal to a patch or across interface curves) and can even be discontinuous across the interface curves. The larger the rate of the mesh size variation, the worse is the shape quality of the resulting mesh. To control this size variation, various methodologies based on metric reduction have been proposed in the case of a continuous isotropic metric. We introduce a novel iterative mesh gradation approach for discontinuous metrics. The approach uses a particular metric reduction procedure in order to ensure the convergence of the gradation process. In particular, we show that in the worst case the anisotropic discontinuous geometric metric map is reduced to an isotropic continuous geometric metric map for which the gradation is controlled [27].

3.12. Linéarisation et maillage des surfaces paramétréées

Participants: P. Laug, H. Borouchaki [correspondant], E. Renaut

Nous proposons une méthodologie pour simplifier la paramétrisation des surfaces composées de carreaux paramétrés issues généralement des environnements CAO (conception assistée par ordinateur). Cette nouvelle paramétrisation est définie via la construction d’une triangulation adaptative appelée support pour chaque carreau. L’adaptation est gouvernée par le contrôle de l’écart entre la paramétrisation initiale et celle issue de la triangulation support. Ce support est utilisé pour générer les maillages de ces surfaces selon une approche indirecte, dans laquelle le maillage est généré via les domaines des paramètres. Le support permet de déconnecter le mailleur du système CAO, et en ce sens constitue un outil universel pour le maillage de telles surfaces [19].

3.13. Mesh adaptation for very high-order numerical scheme

Participants: F. Alauzet, A. Loseille [correspondant] and E. Mbinky

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

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

3.14. Méthodes de remaillages adaptatifs pour le formage incrémental de tôles minces

Participants: L. Moreau [correspondant], A. Cherouat, H. Borouchaki
Développement d’une méthode de remaillage adaptatif surfacique 3D permettant de raffiner de déraffiner le maillage localement autour de l’outil sphérique au cours des simulations numériques de formage incrémental.

3.15. Méthodes de remaillages adaptatifs surfacique dans le cadre des simulations numériques d’emboutissage de structure minces

Participants : L. Moreau [correspondant], A. Cherouat, H. Borouchaki

Développement de méthodes de remaillage adaptatif surfacique 3D, développement de méthodes d’interpolation et transfert des champs, interfaçage avec le code EF Abaqus et application sur des exemples concrets de mise en forme de structure métalliques et composites.

3.16. Multi-scale anisotropic mesh adaptation for unsteady problems

Participants : F. Alauzet [correspondant] and G. Olivier

We focused on the extension of the multi-scale anisotropic mesh adaptation to unsteady flows. It leads to the development of a global fixed point mesh adaptation algorithm and space-time error estimates. Moreover, the mesh adaptation methodology has been extended to the case of moving meshes simulations [25].

3.17. Parallel CAD surface meshing

Participants : P. Laug, H. Borouchaki [correspondant]

A wide range of surfaces can be defined by means of composite parametric surfaces as is the case for most CAD modelers. There are, essentially, two approaches to meshing parametric surfaces: direct and indirect. Popular direct methods include the octree-based method, the advancing-front-based method and the paving-based method working directly in the tridimensional space. The indirect approach consists in meshing the parametric domain and mapping the resulting mesh onto the surface. Using the latter approach, we have proposed a general meshing scheme which consists in discretizing each interface curve and meshing each parametric domain according to the above boundary discretizations. Complex surfaces such as a car engine or a complete aircraft are composed of thousands of patches, and meshing these surfaces using the above sequential scheme can be inefficient. We propose a parallel version of the general meshing scheme while naturally balancing the load to each processor.

3.18. Quelques avancées dans les algorithmes de maillages tétraédriques

Participants : H. Borouchaki, P.-L. George [correspondant], L. Marechal

Le comportement en complexité des algorithmes de triangulation sur les “gros” maillage nous amène à utiliser les algorithmes de renumérotation de type Hilbert qui minimise les défauts de cache. Cette technique est également utilisée comme aide à l’optimisation des “gros” maillages. L’algorithme de renumérotation est multi-œurs.

Des triangulations de plusieurs dizaines de millions de sommets sont construites en utilisant un “simple” ordinateur. La vitesse d’insertion trole le million de tétraèdre à la seconde.

3.19. Reconstruction de surface 3D à partir d’images numériques 2D

Participants : L. Moreau [correspondant], H. Borouchaki, A. Cherouat

Mise au point d’une méthode d’acquisition de la morphologie du buste féminin, développement d’algorithmes de reconstruction de surface 3D d’un buste féminin à partir de photos numériques dans l’objectif de modéliser le comportement du sein en modes statiques et dynamiques.

3.20. Validité des éléments finis usuels

Participants : H. Borouchaki, P.-L. George [correspondant], P. Laug, L. Marechal
On étudie les conditions assurant la validité géométrique des éléments finis usuels de degré 1 et 2. La formulation éléments finis mne conduisant pas toujours à une conclusion simple, on formule les éléments finis sous leur forme de Bézier. Ceci conduit à exhiber des conditions suffisantes (parfois nécessaires et suffisantes) de validité des éléments, c’est-à-dire de la positivité de leur jacobien. Pour les éléments de degré 2, on donne l’interprétation géométrique de ces conditions. Les éléments étudiés sont le triangle à 3 nœuds, le triangle à 6 nœuds, le quadrilatère à 4 nœuds et les quadrilatères à 8 et 9 nœuds, le tétraèdre à 4 nœuds et le tétraèdre à 10 nœuds puis les hexaèdres à 8, 27 et 20 nœuds.

Le cas du simplexe de dimension quelconque et de degré quelconque est traité.

Quelques remarques sur les maillages de surface de degré 2 indiquent quelques pistes à suivre ou à éviter lors de la construction de tels éléments.

### 3.21. Visualisation et modification des maillages courbes d’ordre élevé

**Participants** : J. Castelneau, A. Loseille [correspondant], L. Maréchal

Dans le cadre du projet IIab, des nouveaux algorithmes de visualisation et de modifications interactives des maillages courbes et hybrides ont été développés. En effet, une des principales difficultés dans la génération de maillages courbes reste la visualisation. Il est également nécessaire de disposer d’algorithmes de corrections interactifs car les maillages de surfaces initiaux (de degré 2) sont pour la plupart faux.
5. New Results

5.1. Numerical probability

5.1.1. Simulation of stochastic processes


5.1.1.1. Pathwise convergence of the Euler scheme

A. Alfonsi, B. Jourdain and A. Kohatsu-Higa are studying the convergence for the Wasserstein distance of the Euler scheme towards the limit diffusion. This would be a way to analyze the weak pathwise convergence of this discretization scheme. They have obtained some promising results in the one-dimensional case.

5.1.2. Multi-dimensional models and correlation issues

A. Alfonsi and his PhD Student A. Ahdida have submitted a paper on the simulation of Wishart processes. They have introduced a new family of stochastic differential equations that are defined in the space of correlation matrices and provided high order discretization schemes for such processes [51]. They are currently trying to use this type of matrix valued processes to model the dependence between assets. In particular, they would like to calibrate Index options data. This work is still in progress. The thesis of A. Ahdida has been defended on December 1st [12].

The PhD student Lokmane Abbas Turki has worked on numerical methods for American option pricing based on Malliavin calculus and parallel implementation. He has submitted a paper (co-authored with B. Lapeyre) [50]. He is now working on the dependence of option prices with respect to correlations between stocks in multi-dimensional models.

A. Alfonsi and S. De Marco (postdoc CERMICS) have studied how some option prices (such as spread options) are modified when the correlation between stocks is increased.

5.1.3. Stochastic volatility models

Exotic options and stochastic volatility models is the subject of Sidi Mohamed OULD ALY’s thesis, defended on June 16th, 2011. Sidi-Mohamed has results on the effective computation of option prices in a stochastic volatility model, in the context of variance swap modelling. He has worked out a new model, in the spirit of Bergomi’s approach. This model has remarkable features in terms of tractability and calibration. S.M. Ould Aly has developed numerical methods and an original variance reduction method for models with a log normal volatility. He also has results on the monotony of option prices with respect to the correlation between the stock price and the volatility in the Heston model. Sidi Mohamed has submitted three papers.

5.2. American options

Participants: Benjamin Jourdain, Maxence Jeunesse, Damien Lamberton, Ayech Bouselmi.

5.2.1. American put option with discrete dividends

B. Jourdain and M. Jeunesse are interested in the regularity of the optimal exercise boundary for the American Put option when the underlying asset pays discrete dividends at known times during the lifetime of the option. The dividend amounts are deterministic functions of the asset prices just before the dividend dates. B. Jourdain and M. Jeunesse have proved continuity of the exercise boundary and smooth contact for the value function under general assumptions on the dividend functions.
5.2.2. American options in exponential Lévy models

D. Lamberton and his PhD student A. Bouselmi are working on American options within multi-dimensional exponential Lévy models. They also have preliminary results on the asymptotic behaviour of the exercise boundary of the American put near maturity in the one dimensional case when the limit is strictly smaller than the strike price.

5.3. Financial issues modelling

Participants: Aurélien Alfonsi, Céline Labart, Jérôme Lelong, J. Acevedo.

5.3.1. Credit risk modelling

A. Alfonsi, C. Labart and J. lelong are studying loss models called “stochastic local intensity models” that have been proposed in the literature. First, they are interested in proving mathematically that these models are well posed (it exists and has a unique solution). Second, they aim to provide numerical tools to sample such dynamics.

5.3.2. Limit order markets

A. Alfonsi has an active collaboration with A. Schied (Mannheim University) on limit order book models. A. Schied has visited the CERMICS two weeks in February 2011 and one week on December. They are currently studying some type of non-Markovian resilience for the limit-order book for which they are able to get the optimal execution strategy in a closed form. Moreover A. Alfonsi and his PhD student J. Acevedo study impact models for which the limit order book shape evolves along the time.

5.3.3. Control of systemic risk

Participants: Agnès Sulem, J.Ph. Chancelier, Andreea Minca.

We are interested in contagion modeling and systemic risk in financial networks. We aim to contribute in particular to the domain of control of such systems in order to reduce the systemic risk. We model the propagation of distress in financial systems as an epidemic on a random graph in which the nodes represent financial institutions and edges the exposure between them. Cascade dynamics may be reduced to the evolution of a multi-dimensional markov chain that corresponds to a sequential discovery of exposures and determines the size of contagion. We study the optimal intervention strategy by a lender of last resort who wants to minimize the size of contagion under budget constraints.

5.4. Stochastic control of jump processes and stochastic differential games

Participants: Agnès Sulem [in collaboration with B. Øksendal (Oslo University) and T. Zhang (Manchester University)], John Joseph Absalom Hosking.

5.4.1. Stochastic control under model uncertainty

In [58], we study optimal stochastic control problems with jumps under model uncertainty. We rewrite such problems as (zero-sum) stochastic differential games of forward-backward stochastic differential equations. We prove general stochastic maximum principles for such games, both in the zero-sum case (finding conditions for saddle points) and for the non-zero sum games (finding conditions for Nash equilibria). We then apply these results to study optimal portfolio and consumption problems under model uncertainty. We combine the optimality conditions given by the stochastic maximum principles with Malliavin calculus to obtain a set of equations which determine the optimal strategies.

In [45], we consider some robust optimal portfolio problems for markets modeled by (possibly non-Markovian) Itô–Lévy processes. Mathematically the situation can be described as a stochastic differential game, where one of the players (the agent) is trying to find the portfolio which maximizes the utility of her terminal wealth, while the other player (“the market”) is controlling some of the unknown parameters of the market (e.g. the underlying probability measure, representing a model uncertainty problem) and is trying to minimize this maximal utility of the agent. This leads to a worst case scenario control problem for the agent.
In the Markovian case such problems can be studied using the Hamilton-Jacobi-Bellman-Isaacs (HJBI) equation, but these methods do not work in the non-Markovian case. We approach the problem by transforming it to a stochastic differential game for backward stochastic differential equations (BSDE game). Using comparison theorems for BSDEs with jumps we arrive at criteria for the solution of such games, in the form of a kind of non-Markovian analogue of the HJBI equation. The results are illustrated by examples.

### 5.4.2. Singular stochastic control

In [59], A. Sulem and B. Øksendal study partial information, possibly non-Markovian, singular stochastic control of Itô–Lévy processes and obtain general maximum principles. The results are used to find connections between singular stochastic control, reflected BSDEs and optimal stopping in the partial information case. As an application we give an explicit solution to a class of optimal stopping problems with finite horizon and partial information. Singular control of SPDEs.

In [57], A. Sulem, B. Øksendal and T. Zhang study general singular control problems for random fields given by a stochastic partial differential equation (SPDE). They show that under some conditions the optimal singular control can be identified with the solution of a coupled system of SPDE and a kind of reflected backward SPDE (RBSPDE). They also establish existence and uniqueness of solutions of RBSPDEs.

### 5.4.3. Optimal control with delay

In [44], we study optimal control problems for (time-)delayed stochastic differential equations with jumps. We establish sufficient and necessary stochastic stochastic maximum principles for an optimal control of such systems. The associated adjoint processes are shown to satisfy a (time-) advanced backward stochastic differential equation (ABSDE). Several results on existence and uniqueness of such ABSDEs are shown. The results are illustrated by an application to optimal consumption from a cash flow with delay.

In [48], we will prove a sufficient necessary stochastic maximum principles for the optimal control of SPDEs with delay and study associated time-advanced backward stochastic partial differential equations.

### 5.4.4. Stochastic differential games

In [55], J. Hosking has constructed a stochastic maximum principle (SMP) which provides necessary conditions for the existence of Nash equilibria in a certain form of N-agent stochastic differential game (SDG) of a mean-field type. The information structure considered for the SDG is of a possible asymmetric and partial type. To prove our SMP we use a spike-variation approach with adjoint representation techniques, analogous to that of S. Peng in the optimal stochastic control context. In our proof we apply adjoint representation procedures at three points. The first-order adjoint processes are defined as solutions to certain mean-field backward stochastic differential equations (BSDEs), and second-order adjoint processes of a fist type are defined as solutions to certain BSDEs. Second order adjoint processes of a second type are defined as solutions of backward stochastic equations of a type that we introduce in this paper, and which we term conditional mean-field BSDEs. From the resulting representations, we show that the terms relating to these second-order adjoint processes of the second type are of an order such that they do not appear in our final SMP equations.

### 5.5. Risk measures, BSDEs with jumps and nonlinear expectations

**Participants:** Agnès Sulem, Marie-Claire Quenez, Z. Chen [Shandong University].

In the Brownian case, links between dynamic risk measures and Backward Stochastic Differential Equations (BSDEs) have been established. A. Sulem and M.-C. Quenez are exploring these links in the case of stochastic processes with jumps. They have extended some comparison theorems for BSDEs with jumps, and provided a representation theorem of convex dynamic risk measures induced by BSDEs with jumps. They study optimal stopping problems for (non necessarily) convex dynamic risk measures induced by BSDEs with jumps and establish their connections with Reflected BSDEs with jumps. They also study the case of model ambiguity and its relation with mixed control/optimal stopping problems.
There are two classes of nonlinear expectations, one is the Choquet expectation given by Choquet (1955), the other is the Peng’s $g$-expectation given by Peng (1997) via backward differential equations (BSDE). Recently, Peng raised the following question: can a $g$-expectation be represented by a Choquet expectation? In [ 26 ], A. Sulem and Z. Chen provide a necessary and sufficient condition on $g$-expectations under which Peng’s $g$-expectation can be represented by a Choquet expectation for some random variables (Markov processes). It is well known that Choquet expectation and $g$-expectation (also BSDE) have been used extensively in the pricing of options in finance and insurance. Our result also addresses the following open question: given a BSDE ($g$-expectation), is there a Choquet expectation operator such that both BSDE pricing and Choquet pricing coincide for all European options? Furthermore, the famous Feynman-Kac formula shows that the solutions of a class of (linear) partial differential equations (PDE) can be represented by (linear) mathematical expectations. As an application of our result, we obtain a necessary and sufficient condition under which the solutions of a class of nonlinear PDE can be represented by nonlinear Choquet expectations [ 26 ].

5.6. Malliavin calculus and applications

5.6.1. Lower bounds for the density of functionals on the Wiener space

In collaboration with: B. Fernandez and A. Meda from the University of Mexico, V. Bally gave a lower bound for general Itô processes to remain in a tube around a given curve. This is done under some ellipticity assumption in [ 21 ]. Now, with L. Caramellino (University Tor Vergata, Rome) he investigates the case of a diffusion processes which satisfies the Hörmander condition.

5.6.2. Malliavin Calculus for Poisson Point Processes and applications

V. Bally and E. Clément (Université Paris-Est Marne la Vallée) study the density of the law of the solution of a stochastic equation with jumps, which has discontinuous coefficients [ 18 ], [ 19 ]. Moreover, with N. Fournier (university of Creteil), V. Bally obtained results on the smoothness of the law of a bidimensional Boltzman equation [ 22 ].

5.6.3. Riesz transform and regularity of the law of a random variable

The idea of using the Riesz transform in order to study the regularity of the law of a random variable appears in former works of P. Malliavin and A. Thalmaier. In collaboration with L. Caramellino (University Tor Vergata, Rome) we gave regularity results using this tool [ 17 ].
5. New Results

5.1. Computational quantum chemistry

Participants: Eric Cancès, Ismaila Dabo, Virginie Ehrlacher, Salma Lahbabi, Francis Nier, Gabriel Stoltz.

In computational quantum chemistry 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, V. Ehrlacher, S. Lahbabi and G. Stoltz have addressed issues related to the modeling and simulation of local defects in periodic crystals (see [61] for a pedagogical introduction).

Computing the energies of local defects in crystals is a major issue in quantum chemistry, materials science and nano-electronics. In collaboration with M. Lewin (CNRS, Cergy), E. Cancès and A. Deleurence have proposed in 2008 a new model for describing the electronic structure of a crystal in the presence of a local defect. This model is based on formal analogies between the Fermi sea of a perturbed crystal and the Dirac sea in Quantum Electrodynamics (QED) in the presence of an external electrostatic field. The justification of this model is obtained using a thermodynamic limit of Kohn-Sham type models. In [29], E. Cancès and G. Stoltz studied the time evolution of defects within this model, in the context of linear response, which allowed them to give a rigorous meaning to the Adler-Wiser formula for the frequency-dependent dielectric permittivity of crystals. In [27], E. Cancès et V. Ehrlacher have proved that local defects are always neutral in the Thomas-Fermi-von Weizsäcker (TFW) theory. In this respect, all TFW crystals behave like metals. In collaboration with M. Lewin, E. Cancès and S. Lahbabi are working on the extension of Kohn-Sham like models to disordered systems.

On the numerical side, E. Cancès has worked with Y. Maday and R. Chakir (University Paris 6) on the numerical analysis of the electronic structure models. In [26], they have obtained optimal \(a_{\text{priori}}\) error bounds for the the planewave approximation of the Thomas-Fermi-von Weizsäcker and the Kohn-Sham LDA models. Together with Y. Maday, E. Cancès and V. Ehrlacher have analyzed the computation of eigenvalues in spectral gaps of locally perturbed periodic Schrödinger operators [28].

Photovoltaic cells based upon organic photovoltaic (OPV) semiconductors are cost-effective, light-weight alternatives to conventional silicon and thin-film technologies. This year, in collaboration with A. Ferretti, N. Poilvert, N. Marzari (MIT and University of Oxford), M. Cococcioni (University of Minnesota), and Y. L. Li (Xiamen University), I. Dabo has worked on the prediction of the electronic spectra and electrical response of molecular and polymer compounds for OPV applications. Specifically, Y. L. Li and I. Dabo have implemented an efficient computational method for studying the electrical response of semiconducting polymers, in close agreement with more expensive local basis-set algorithms [52]. In parallel, A. Ferretti, I. Dabo, M. Cococcioni, and N. Marzari have applied recently developed electronic-structure theories, namely, orbital-dependent density-functional theories (OD-DFTs), to describe donor and acceptor levels in semiconducting organic materials, demonstrating that OD-DFT is apt at describing donor and acceptor levels within 0.1-0.4 and 0.2-0.6 eV of experiment [41]. Future work includes the prediction of the band structure of semiconductor alloys and the simulation of molecular heterojunctions.

The current works by F. Nier related to electronic structure calculations concern the mean field limit of the quantum dynamics in the bosonic setting; and the non linear modelling of Bose-Einstein condensates. A few years ago, Z. Ammari and F. Nier initiated a program about the Hamilton mean field dynamics of a large numbers of bosons. Their approach reconsiders the old idea that the mean field limit in the bosonic setting is actually a semiclassical limit in infinite dimension. In previous works, they proved results which are general with respect to the initial data but hold for bounded interactions. They also proved that the dynamics of the BBGKY hierarchy, often used within the mean field theory, is actually a projected picture of the dynamics of Wigner measures in the infinite dimensional phase-space. Recently in [10], they obtained such results for
singular pair interaction potentials, which include the Coulombic (attractive or repulsive) case in dimension $d = 3$. Regarding the modelling of Bose-Einstein condensates, F. Nier and collaborators have developed in [9] complete analysis of the nonlinear adiabatic ansatz proposed by physicists for the simulation of rotating Bose-Einstein condensates.

### 5.2. Computational Statistical Physics

**Participants:** Matthew Dobson, Claude Le Bris, Frédéric Legoll, Tony Lelièvre, Francis Nier, Stefano Olla, Grigorios Pavliotis, Giovanni Samaey, Gabriel Stoltz.

The extremely broad field of Molecular dynamics (MD) is a domain where the MICMAC project-team, originally more involved in the quantum chemistry side, has invested a lot of efforts in the recent years. Molecular dynamics may also be termed computational statistical physics since the main aim is to numerically estimate average properties of materials as given by the laws of statistical physics. The project-team studies both deterministic and probabilistic techniques used in the field.

#### 5.2.1. Free Energy calculations

For large molecular systems, the information of the whole configuration space may be summarized in a few coordinates of interest, called reaction coordinates. An important problem in chemistry or biology is to compute the effective energy felt by those reaction coordinates, called free energy.

In the article [51], Tony Lelièvre, Mathias Rousset and Gabriel Stoltz study the application of constrained Langevin dynamics to the computation of free energy differences, by thermodynamic integration techniques and fluctuation relation (à la Jarzynski).

The work by T. Lelièvre and K. Minoukadeh on the longtime convergence of the ABF method in a particular bi-channel scenario (which was already mentioned in last year’s activity report) has been accepted for publication [50]. Likewise, the work by Nicolas Chopin (CREST, ENSAE), T. Lelièvre and G. Stoltz on application of the ABF method to Bayesian inference is about to appear, see [30].

#### 5.2.2. Sampling trajectories

There exist a lot of methods to sample efficiently Boltzmann-Gibbs distributions. The situation is much more intricate as far as the sampling of trajectories (and especially metastable trajectories) is concerned.

In [32], T. Lelièvre and D. Pommier, in collaboration with F. Cérou and A. Guyader (INRIA Rennes, ASPI) investigated the interest of an Adaptive Multilevel Splitting algorithm to compute reactive paths, and estimate transition rates. The obtained results are very interesting. Current research aims at testing the technique on practical cases.

In [49], C. Le Bris and T. Lelièvre together with M. Luskin and D. Perez have proposed a mathematical analysis of the parallel replica algorithm, introduced by A. Voter in 1997 to efficiently simulate metastable trajectories. This work opens a lot of perspectives using a generic tool (the quasi stationary distribution) to relate continuous state space dynamics (Langevin type dynamics) to discrete state space dynamics (kinetic Monte Carlo type models). A follow-up work consists in theoretically investigating another related approach, the hyperdynamics method.

#### 5.2.3. Nonequilibrium systems

The efficient simulation of molecular systems is known to be a much more complicated problem when the system is subjected to a non-conservative external forcing than when the system experiences conservative forces. Together with the sampling of metastable dynamics mentioned above, these are the two major research focus in MD of the project-team.
On this topic, G. Stoltz continued his long lasting collaboration with physicists at CEA/DAM on reduced models for shock and detonation waves. More precisely, he published two works applying some simulation techniques he devised to actual energetic materials of interest to physicists, namely (i) a technique to sample constraints in average and allowing to compute the Hugoniot curve efficiently, which was applied to reacting TATB [24]; and (ii) a reduced stochastic dynamics to model detonation waves, applied to a material with properties close to nitromethane and allowing an atomistic simulation of the shock-to-detonation transition [55].

F. Legoll and G. Stoltz pursued their studies of the anomalous thermal conductivity of one dimensional chains. They have investigated the case of a chain of rotors subjected to a mechanical forcing. In collaboration with A. Iacobucci and S. Olla (CEREMADE, Paris Dauphine) they have shown in [44] that the mechanical forcing can have a counter-intuitive effect and reduce the thermal current. Besides this numerical study, G. Stoltz, in collaboration with C. Bernardin (ENS Lyon) considered in [18] the issue of thermal transport in one of the simplest possible one dimensional model, a chain of oscillators whose kinetic and potential energy functions are the same, and which are subjected to a stochastic noise exchanging all the variables. The system therefore has only two conservation laws, the energy and the total length. A hydrodynamic limit consisting of a system of conservation laws can be obtained before the onset of shocks. However, the thermal transport is anomalous: this can be proved by analytical computations for harmonic interactions, or demonstrated numerically in the general case.

G. Stoltz also studied techniques to compute the viscosity of fluids using steady state nonequilibrium dynamics with an external nongradient bulk forcing, in the framework of the context of the PhD of Rémi Joubaud, see [45]. The two authors have proved a linear response result, and obtained asymptotic scalings of the viscosity in terms of the friction coefficients of the underlying Langevin dynamics. G. Stoltz and G. Pavliotis are now extending the results to the case of time dependent nongradient external forcings.

Nonequilibrium molecular dynamics simulations can also be used to compute the constitutive relation between the strain rate and stress tensor in complex fluids. This is fulfilled simulating molecular systems subject to a steady, non-zero macroscopic flow at a given temperature. Starting from a bath model, M. Dobson, F. Legoll, T. Lelievre, and G. Stoltz have derived a Langevin-type dynamics for a heavy particle in a non-zero background flow. The resulting dynamics, which is theoretically obtained when a unique large particle is considered, is numerically observed to also perform well when a system of many interacting particles within shear flow is considered.

5.2.4. Effective dynamics

For a given molecular system, and a given reaction coordinate $\xi : \mathbb{R}^n \rightarrow \mathbb{R}$, the free energy completely describes the statistics of $\xi(X)$ when $X \in \mathbb{R}^n$ is distributed according to the Gibbs measure. On the other hand, obtaining a correct description of the dynamics along $\xi$ is complicated.

F. Legoll and T. Lelièvre have continued their work on the definition and the analysis of a coarse-grained dynamics that approximates $\xi(X_t)$, when the state of the system $X_t$ evolves according to the overdamped Langevin equation (which is ergodic for the Gibbs measure). The aim is to get a coarse-grained description giving access to some dynamical quantities, such as residence times in metastable basins. These basins are usually assumed to be completely described by $\xi$. They have proposed an effective dynamics, which is derived using conditional expectations. The first accuracy result, obtained in 2010, is phrased in terms of an estimate on the relative entropy between the law of $\xi(X_t)$ and the law of its approximation, at any time $t$ (see [63] for a description of the results in a simple case). If an appropriate time-scale separation is present in the system, then the effective dynamics is accurate in the sense of time-marginals. The obtained numerical results showed that this dynamics can also be used to accurately compute residence times in potential energy wells, and thus seem to be accurate in a much stronger sense. Together with S. Olla, they have started to analyze the pathwise accuracy of the proposed coarse-grained dynamics.

The extension of the numerical strategy to the case when the reference dynamics on the whole system is the Langevin dynamics has also been studied. Promising numerical results have already been obtained in collaboration with a master’s student (F. Galante).
5.2.5. Convergence to equilibrium

An important question for the analysis of sampling techniques is the rate of convergence to equilibrium for stochastic trajectories.

F. Nier continues to investigate the spectral properties of Witten Laplacians and Kramers-Fokker-Planck operators. In a recent collaboration with D. Le Peutrec and C. Viterbo, the Arrhenius law for metastable states, in its refined version also known as Eyring-Kramers law, was extended to $p$-forms. Very accurate analytical results have been provided for the exponentially small eigenvalues of Witten Laplacians acting on $p$-forms.

F. Nier has started a collaboration with T. Lelièvre about accurate exit laws for Smoluchowski processes, via a Witten complex approach. Investigations with G. Pavliotis and T. Lelièvre have started about non gradient diffusion systems.

5.2.6. Hamiltonian dynamics

Constant energy averages are often computed as long time limits of time averages along a typical trajectory of the Hamiltonian dynamics. One difficulty of such a computation is the presence of several time scales in the dynamics: the frequencies of some motions are very high (e.g. for the atomic bond vibrations), while those of other motions are much smaller. This problem has been addressed in a two-fold manner.

Fast phenomena are often only relevant through their mean effect on the slow phenomena, and their precise description is not needed. Consequently, there is a need for time integration algorithms that take into account these fast phenomena only in an averaged way, and for which the time step is not restricted by the highest frequencies. In [38], M. Dobson, C. Le Bris, and F. Legoll developed integrators for Hamiltonian systems with high frequencies. The integrators were derived using homogenization techniques applied to the Hamilton-Jacobi PDE associated to the Hamiltonian ODE. This work extends previous works of the team. The proposed algorithms can now handle the case when the (unique) fast frequency depends on the slow degrees of freedom, or when there are several fast constant frequencies.

Another track to simulate the system for longer times is to resort to parallel computations. An algorithm in that vein is the parareal in time algorithm. It is based on a decomposition of the time interval into subintervals, and on a predictor-corrector strategy, where the propagations over each subinterval for the corrector stage are concurrently performed on the processors. Using a symmetrization procedure and/or a (possibly also symmetric) projection step, C. Le Bris and F. Legoll, in collaboration with X. Dai and Y. Maday, have introduced several variants of the original plain parareal in time algorithm [35]. These variants, compatible with the geometric structure of the exact dynamics, are better adapted to the Hamiltonian context.

5.3. Complex fluids

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

In [48], Claude Le Bris and Tony Lelièvre have reviewed the state-of-the-art of numerical and mathematical results on micro-macro models for viscoelastic fluids.

Following previous works, Claude Le Bris and Tony Lelièvre together with Lingbing He have analyzed in [43] the longtime behaviour of nematic polymeric fluids (liquid crystals). The longtime asymptotics for such models is much richer than for flexible polymers, which were considered in previous analysis. Indeed, for these models, periodic in time behaviours are expected.

In his PhD under the supervision of Claude Le Bris and Tony Lelièvre, David Benoit studies models of aging fluids developed at the ESPCI (Ecole supérieure de physique et de chimie industrielles) and designed to take into account phenomena such as shear thinning, aging and shear banding in falling sphere experiments. The work consists on the one hand in studying the mathematical well-posedness of some macroscopic models and on the other hand in trying to understand the link between such macroscopic models and microscopic models which have been proposed to describe such fluids.
In the line of his former work [15], Sébastien Boyaval has pursued his work about the mathematical modelling of viscoelastic fluid flows. A new reduced model for thin layers of elastic gravity flows with a free surface was derived in collaboration with François Bouchut, following similar hydrostatic assumptions to those which lead to the Saint-Venant equations as a usual reduced model for shallow water flows. The model is naturally endowed with an energy but the conservative part is non-standard: the energy is not convex with respect to the conservative variables. It is convex with respect to other (more physical) variables. For the numerical simulation of possibly discontinuous solutions, a relaxation scheme is proposed in order to ensure that the numerical approximation mimics the natural energy dissipation.

In [58], Tony Lelièvre together with Giovanni Samaey and Vincent Legat explored some numerical techniques to get closed macroscopic equations from microscopic models. The proposed method can be seen as a way to justify and extend techniques based on the so-called quasi-equilibrium approximation.

In [40], the effect of a topography on a free surface flow is studied using the free-surface Navier Stokes equations and ALE method for discretization.

5.4. Application of greedy algorithms

Participants: Sébastien Boyaval, Eric Cancès, Virginie Ehrlacher, Tony Lelièvre.

The manuscript [25] where E. Cancès, V. Ehrlacher and T. Lelièvre study the convergence of greedy algorithms to nonlinear convex problems has been accepted for publication. Current research now aims at extending such techniques to non-symmetric problems.

S. Boyaval has continued his study of a new variance reduction method introduced with T. Lelièvre. It makes an innovative use of a greedy algorithm in order to construct control variates for a parameterized family of Monte-Carlo estimator. New successful applications are typical problems in uncertainty quantification.

5.5. Homogenization

Participants: Ronan Costaouec, Claude Le Bris, Frédéric Legoll, Francis Nier, Florian Thomines.

The project-team has pursued its efforts in the field of stochastic homogenization of elliptic equations. The various contributions of the team, which aim at designing numerical approaches that both are practically relevant and keep the computational workload limited, have been presented from a unified perspective in [59].

An interesting case in that context is when the randomness comes as a small perturbation of the deterministic case. As previously shown by former works of the project-team, this situation can indeed be handled with a dedicated approach, which turns out to be far more efficient than the standard approach of stochastic homogenization. This previous analysis was performed manipulating the exact correctors, solutions to PDEs posed on $\mathbb{R}^d$. In practice, one can only consider a truncated version of these corrector problems, which is next discretized using e.g. a Finite Element method. The previous analysis has been extended to this practical situation by R. Costaouec [31].

In the work mentioned above, the perturbation to the deterministic case is supposed to be small in the $L^\infty$ norm (that is, almost surely small). The team has also considered the case when the perturbation is small in a weaker norm, typically a $L^p$ norm with $p < \infty$ (the case when only the expectation of the perturbation is assumed to be small, rather than the perturbation itself, is covered by that framework). The approach proves to be very efficient from a computational viewpoint, in comparison to the standard approach of stochastic homogenization, as shown in [13], [12]. In that setting, the computation of the homogenized matrix requires repeatedly solving a corrector-like equation for various configurations of the material. For this purpose, C. Le Bris and F. Thomines have shown how to adapt the Reduced Basis approach to the specific context, to even further reduce the computational cost [62].
The team has also proceeded to address, from a numerical viewpoint, the case when the randomness is not small. In that case, using the standard homogenization theory, one knows that the homogenized tensor, which is a deterministic matrix, depends on the solution of a stochastic equation, the so-called corrector problem, which is posed on the whole space $\mathbb{R}^d$. This equation is therefore delicate and expensive to solve. In practice, the space $\mathbb{R}^d$ is truncated to some bounded domain, on which the corrector problem is numerically solved. In turn, this yields a converging approximation of the homogenized tensor, which happens to be a random matrix. For a given truncation of $\mathbb{R}^d$, R. Costaouec, C. Le Bris and F. Legoll, in collaboration with X. Blanc (CEA), have studied how to reduce the variance of this matrix, using the technique of antithetic variables, which is a method widely used in other fields of application. Its efficiency in the context of stochastic homogenization has been extensively studied, both numerically and theoretically [60], [19]. R. Costaouec, C. Le Bris and F. Legoll are currently investigating the possibility to use other variance reduction approaches, such as control variate techniques.

From a numerical perspective, the Multiscale Finite Element Method is a classical strategy to address the situation when the homogenized problem is not known, or when the scale of the heterogeneities, although small, is not considered to be zero (and hence the homogenized problem cannot be considered as a sufficiently accurate model). The extension of this strategy to the stochastic case, when the tensor describing the properties of the material is the sum of a periodic term and a small random term, has been studied by C. Le Bris, F. Legoll and F. Thomines [46]. A method with a much smaller computational cost than the original MsFEM in the stochastic setting has been proposed. Provided the stochastic perturbation is indeed small, the proposed method is as accurate as the original one. The work [46] also provides a complete analysis of the approach, extending that available for the deterministic setting. Such an analysis often relies on the rate of convergence of the two scale expansion (in the sense of homogenization theory) of the solution to the highly oscillatory elliptic partial differential equation. The result is classic for periodic homogenization. In generic stochastic homogenization, the rate can be arbitrary small, depending on the rate with which the correlations of the random coefficient vanish. In [47], such a result has been established for weakly stochastic homogenization. This result is a key ingredient for the numerical analysis of the MsFEM approach proposed in [46].

Still in the framework of the Multiscale Finite Element approach, F. Thomines has further investigated, in collaboration with Y. Efendiev and J. Galvis (Texas A&M University), the use of Reduced Basis methods. They have considered an extension of the MsFEM approach, well suited to the high contrast case, i.e. the case when the ratio between the maximum and the minimum values of the heterogeneous coefficient is large. The main idea of this extension is to complement the standard MsFEM basis functions with the eigenfunctions (associated to the first small eigenvalues) of a local eigenproblem. In [39], Y. Efendiev, J. Galvis and F. Thomines have considered the case when the problem depends on an additional parameter, and shown how to use the Reduced Basis approach to more efficiently compute the eigenfunctions mentioned above.

The theoretical results obtained by the team on variance reduction [19] and on the rate of convergence of the two scale expansion of the solution to a highly oscillatory, weakly random PDE [47], both rely on asymptotic properties of the Green function of the elliptic operator $Lu = -\text{div} (A \nabla u)$, where $A$ is a periodic, coercive and bounded matrix. In collaboration with X. Blanc (CEA) and A. Anantharaman, F. Legoll has established some results of this question [11]. This contribution presents in a unified manner and complements several results already given in the literature.

All the works previously mentioned are concerned with elliptic PDEs. F. Nier has studied various problems in the context of wave propagation in random heterogeneous media. In collaboration with S. Breteaux (PhD student in Rennes), he has derived a Boltzmann type equation from first principles of quantum mechanics, using the bosonic QFT presentation of Gaussian random fields. Various extensions of this result are currently under investigation.

### 5.6. Atomistic to continuum methods

**Participants:** Matthew Dobson, Claude Le Bris, Frédéric Legoll.
The project-team has continued their theoretical and numerical efforts on the general topic of "passage from the atomistic to the continuum". This concerns theoretical issues arising in this passage but also the development and the improvement of numerical simulations coupling the two scales.

The quasicontinuum method couples an atomistic model to a continuum approximation in order to compute deformed states of a crystalline lattice at a reduced computational cost compared to a full atomistic simulation. In collaboration with M. Luskin (University of Minnesota) and C. Ortner (Warwick), M. Dobson analyzed the use of numerical solvers for approximating solutions to the equilibrium equations of the force-based quasicontinuum method [37]. In particular, it was shown that a previously-proposed modified conjugate gradient algorithm has unstable modes since the linearized operator is generally not positive-definite. Based on observed properties of the spectrum, convergence rates are given for the GMRES method applied to the operator.

Still in the framework of quasicontinuum methods, several consistent couplings have been proposed in the literature in the past years. M. Dobson showed the impossibility of constructing higher-order consistent couplings for quasicontinuum energies [36]. The analysis is performed in the one-dimensional situation, and is based on the fact that the truncation error gives lower bounds on the global error on the deformation gradient. A consequence of this result is that the so-called quasi-nonlocal energy (which is one of the coupling schemes proposed earlier) has asymptotically optimal error bounds.

The team has also addressed questions related to the finite temperature modeling of atomistic systems and derivation of coarse-grained descriptions, such as canonical averages of observables depending only on a few variables. In the one-dimensional setting, an efficient strategy that bypasses the simulation of the whole system had been proposed in 2010. In collaboration with X. Blanc (CEA), F. Legoll has extended this strategy to the so-called membrane setting [20]: the system is composed of atoms that lie on a two-dimensional lattice, and have a unique degree of freedom, representing their height. The strategy can also be used to derive the stress-strain relation for one-dimensional chains of atoms, e.g. the relation between the elongation of the chain and the stress, at any given temperature [63].

5.7. Surface chemistry and electrochemistry

Participant: Ismaila Dabo.

Rising energy imperatives have revived strong interest in electrochemical cells (e.g., fuel cells) and photoelectrochemical cells (e.g., dye-sensitized solar cells). Understanding and improving the electrical performance of such systems entails the accurate description of electrode-electrolyte interfaces at the microscopic level. Despite recent advances in the application of computational methods to study realistic electrode-electrolyte interfaces, capturing the effect of the applied electrical voltage and solvent electrical response remains a central challenge in computational electrochemistry. In order to address this difficulty, a comprehensive model for chemical systems embedded in ionic media has been developed. The model couples a quantum description of the electrode with a continuum representation of the electrolyte.

This year, in collaboration with O. Andreussi (MIT) and N. Marzari (University of Oxford), I. Dabo has worked on improving the predictive accuracy and numerical implementation of the continuum solvation model to describe molecular systems and metallic surfaces in interaction with a solvent [14]. The revised model overcomes some of the numerical limitations encountered in existing solvation models and extends their range of applicability. The approach proceeds by recasting the problem in terms of induced polarization charges that act as a direct mapping of the continuum dielectric solvent. The model is defined in a self-consistent manner in terms of the first-principles electronic density of the solute, thereby limiting the number of numerical parameters involved in existing solvation methods. The model accounts for additional pressure and cavitation contributions. The resulting self-consistent continuum solvation (SCCS) model provides an effective, compact fit of computational and experimental data, with solvation energies in error of 0.3-0.4 kcal/mol. The model is implemented in the widely used open-source program QUANTUM-ESPRESSO, exploiting a numerical approach that is intrinsically parallel, robust, and straightforward to adapt to most electronic-structure codes.
In terms of practical applications in surface science and electrochemistry, the adsorption of arsenate (a severe contaminant in drinking water) at the surface of biogeochemical minerals has been studied in collaboration with M. Blanchard and G. Morin (Université Paris 6) [22]. The simulation of infrared spectroscopic experiments for carbon monoxide (a severe catalytic poison) at catalytic surfaces has also been addressed in [33]. Future work includes the generalization of the SCCS model to surfaces in contact with a reservoir of electrons, and molecular dynamics simulations of electrode surfaces.
POEMS Project-Team

6. New Results

6.1. Numerical methods for time domain wave propagation

6.1.1. High Order Theta Scheme for the linear wave equation.

Participants: Juliette Chabassier, Sébastien Impériale.

We have pursued our work about a new class of high order implicit three time step schemes for semi-discretized wave equations of the form

\[
\frac{d^2}{dt^2} u_h + K_h u_h = 0, \quad u_h(0) = u_{0,h}, \quad \frac{du_h}{dt}(0) = u_{1,h},
\]

where \( K_h \) is a symmetric positive definite matrix. For such problems, explicit three time step schemes generally show good performances but present two major drawbacks, in some situations, that have not yet been completely overcome:

- If the mesh presents elements of quite different sizes, the time step must be adapted to the smallest one because of the CFL condition.
- If the mass matrix is non diagonal or non block-diagonal, its inversion (at least one time per iteration) can lead to a dramatic over cost of the schemes, which is obviously not the case with implicit schemes for which a matrix has to be inverted in any case.

A natural way to avoid this restriction is to use local time stepping techniques which divides into two categories. First, the locally implicit technique, which is optimal in term of CFL restriction but “only” second order accurate in time, and requires the inversion of interface matrices. Second, the fully explicit local time stepping, as developed that enables to achieve high order time stepping but without (up to now) a full control over the CFL condition.

This is why we construct generalized implicit \( \theta \)-scheme using the modified equation approach to obtain 4th order approximation. These schemes introduce 2 degree of freedom \( \theta \) and \( \varphi \) and can be written under a general form as

\[
\frac{u_h^{n+1} - 2u_h^n + u_h^{n-1}}{\Delta t^2} + K_h \left( \theta u_h^{n+1} + (1 - 2 \theta) u_h^n + \theta u_h^{n-1} \right) + \left( 1 - \frac{\theta}{12} \right) K_h^2 \left( \varphi u_h^{n+1} + (1 - 2 \varphi) u_h^n + \varphi u_h^{n-1} \right) = 0.
\]

The parameters \( \theta \) and \( \varphi \) are chosen as functions of the time step through the resolution of an optimization problem: we consider that the time step is given and we try to obtain a scheme that minimizes the consistency errors still being stable. The limit problem when the time step tends is infinite gives an optimal unconditionally stable fourth order scheme. This work has been submitted for publication.

6.1.2. Discontinuous Galerkin Methods for wave equations

Participants: Patrick Joly, Antoine Tonnoir.
This has been the subject of the internship of A. Tonnoir and can be seen as a contribution to the mathematical analysis of coupled BEM / DG methods for time domain diffraction problems (see section 6.1.4). We did not pretend to make a real contribution to the field of the analysis of DG methods, but wanted to understand in sufficient detail the quite surprising observation that the non-centered (in space) schemes, provided by the use of non centered fluxes in the DG approach, leads to a better accuracy than the centered schemes issued for central fluxes! This has been done (for the simple 1D wave equation) from two points of view: the direct energy method and the Fourier analysis (or dispersion analysis) on regular grids.

6.1.3. Analysis of time domain boundary integral equations

Participants: Patrick Joly, Nicolas Le Guillarme.

This has been developed again as a part of the mathematical analysis of coupled BEM / DG methods for time domain diffraction problems (see section 6.1.4). With J. Rodríguez, we have revisited the analysis of the coercivity / continuity properties of space-time boundary integral associated to wave diffraction problem. Unlike the more traditional approach based on Laplace transform in time (cf the initial work by Bamberger-Ha Duong) we treat this question directly in the time domain by exploiting in a simple way the connection between integral equations and boundary value problems. This can be done in a very general setting but also particularized to the 1D case for which we got sharp estimates (internship of N. Le Guillarme). Furthermore, we can reinterpret the classical discretization by finite elements for retarded potentials as non conforming finite element methods. This allows to investigate in a new way the error analysis of time domain boundary element methods, which will be the subject of a future work.

6.1.4. Coupling Retarded Potentials and Discontinuous Galerkin Methods for time dependent wave propagation problems

Participant: Patrick Joly.

This topic is developed in collaboration with J. Rodríguez (Santiago de Compostela) in the framework of the contract ADNUMO with AIRBUS. Let us recall that our objective was to use time-domain integral equations (developed in particular at IMACS and EADS) as a tool for constructing transparent boundary conditions for wave problems in unbounded media. Our previous contribution of this topic concern the construction of an energy preserving method for the coupling space-time Galerkin approximation of the integral equations with discontinuous Galerkin finite elements and leap-frog time discretization for the numerical approximation of the equations inside the computational domain. For stability reasons, we had to use central fluxes. The drawback of central fluxes is that they do not lead to an optimal accuracy (see also the paragraph on the analysis of DG methods) which is traduced in practice by high frequency spurious oscillations.

Our objective this year was to look for another discretization approach that would overcome these problems. The approach we propose consists in playing on the time discretization procedure. For this we decompose the stiffness bilinear form as the sum of a conservative term corresponding to the use of central fluxes and a ("small") dissipative term due to off-centering and involving jumps across interfaces of the discrete solution. We propose a scheme which treats the conservative part of the equation in a centered way (leap-frog type) and the dissipative term in a non centered way (backward Euler type). Doing so, the overall accuracy in time of the scheme is preserved (with respect to the case of central fluxes) and the stability is maintained at the price of a (slightly) more constraining CFL conditions, which does not seem that much penalizing for the application. The stability is analyzed through the decay of an discrete energy. As a consequence, we can adapt the discretization of the coupling terms between integral and interior equations in order to preserve the stability of the fully coupled scheme under the same CFL condition. Numerical validations are in progress.


Participants: Eliane Bécache, Aliénor Burel, Sébastien Impériale, Patrick Joly.
This topic is the subject of the first part of the PhD thesis of A. Burel and has been developed for a part in collaboration with Marc Duruflé. Decomposing the displacement field into potentials is a well-known tool in elastodynamics, and it expresses the decoupling of the pressure wave and the shear wave inside a homogeneous isotropic media. Although this tool is classically used when searching for analytic solutions, it does not seem to have been exploited for numerical computation using finite elements for instance. However, this is a priori attracting since, contrary to a displacement field approach for instance, it allows to decouple the approximation of P and S waves and to adapt the discretization process (mesh size, order of elements) to the dynamics of each type of wave, which is a priori particularly interesting when S-waves propagate much more slowly than P-waves (soft materials such as rubber). The main difficulty is to cope with the coupling of the different types of waves (the so-called conversion of modes) which occurs, due to wave reflections and transmissions, at interfaces between homogeneous media or at physical boundaries. The simplest situation where this phenomenon appears is the propagation of elastic waves in a homogeneous domain with clamped boundary. The main difficulty is to guarantee the stability of the treatment of the boundary condition. For this, we have proposed a variational formulation in which the stiffness bilinear form appears as a sum of a decoupled volumic bilinear form and a coupled surfacic bilinear form. This formulation is is compatible to an energy conservation result. After space discretization with finite elements spaces well adapted to each type of waves (using different meshes and/or polynomial degrees), we propose a discrete energy preserving numerical scheme, based on an explicit discretization of the volume terms and an implicit discretization of the surface terms. The resulting scheme is mainly explicit (only a sparse boundary linear system has to be solved at each time step) and stable under a CFL stability condition that is not affected by the presence of the boundary. An approach based on a modification of this scheme has been proposed for treating the free surface condition. This approach appears to give satisfactory results in the frequency domain (on the basis of numerical simulations) but fails in tim domain due to (apparently unconditional) instabilities.

6.1.6. Mathematical and numerical modeling of piezoelectric sensors.

Participants: Sébastien Impériale, Patrick Joly.

This research, which constitutes the subject of the PhD thesis of S. Imperiale (which will be defended in January 2012) is developed in the framework of a collaboration with CEA-LIST about the numerical modeling of non-destructive testing experiments using ultrasonic waves. More precisely, we have concerned during the past three years by mathematical and numerical questions related to the simulation of non destructive testing experiments using piezoelectric devices. In particular, we have investigated the modeling of piezoelectric sensors that are used to generate and record ultrasonic waves in a solid material: such waves are typically used to investigate in a non invasive way the possible presence of defects in manufactured items. Such an issue has already been tackled in the engineering literature but not, to our knowledge, by way of rigorous applied mathematics. As we are arriving at the conclusion of this work, let us summarize our main contributions during these three years (these have been published in M2AN):

- The full equations of piezoelectricity couple Maxwell’s equations with linear elastodynamics equations which corresponds to a coupled hyperbolic system. This system presents quite different time scales due to the very large ratio between the speed of light and the sound speed, which makes it impossible to treat by a direct numerical approach. To overcome this problem, we have given a rigorous justification, via asymptotic analysis, of the so-called quasi-static approximation model in which the electric unknowns are reduced to a scalar electric potential: the reduced model appears as a coupled elliptic-hyperbolic system.

\[
\rho \partial_t \mathbf{u} - \nabla \cdot \mathbf{e} \mathbf{\epsilon} - \nabla \cdot \mathbf{e} \mathbf{\varphi} = 0 \quad \text{in } \Omega_S \quad (\text{the solid domain}),
\]
\[
\nabla \cdot \mathbf{e} \mathbf{\varphi} - \nabla \cdot \mathbf{e} \mathbf{\epsilon} (\mathbf{u}) = 0 \quad \text{in } \mathbb{R}^3.
\]

- We have next justified the reduction of the computation of this electric potential to the piezoelectric parts of the computational domain, considering the large contrast of permittivity between piezoelectric materials and surrounding polymers. This relies again on a limit process: the ratio between permittivities is the small parameter.
A particular attention has to be devoted to the modeling of the electric supply process: more precisely the nonlocal (in space) boundary conditions on the electrodes (used to model the emission and reception regimes), as well as the modeling of the coaxial cable connecting the sensor to the electric generator (see also section 6.4.1).

Concerning the numerical approximation, an energy preserving finite element / finite differences numerical scheme is developed. A specific procedure is used for treating the nonlocal boundary conditions on the electrodes. Unbounded media have been treated via PML techniques that are dealt with using an original mortar element technique (see section 6.3.1 for more details). Various numerical results in academic or more realistic situations have been obtained. For instance, we have been able to model how one can exploit multi-component devices (see figure 1) to produce well focused waves (see figure 2).

A computational code issued from our research, named Ondomatic (12 000 lines in C++), has been implemented. This code uses the finite element library FEMME (15 000 lines in C++) that essentially relies upon the use of spectral finite element on hexahedral meshes.

Figure 1.

Figure 2.
6.1.7. Maxwell’s equations in Lorentz materials

**Participants:** Patrick Ciarlet, Patrick Joly, Valentin Vinoles.

This is the time-domain counterpart of the research done at Poems about metamaterials (see also the section 6.2.7) and has been the subject of the internship of V. Vinoles. Lorentz materials are particular non dissipative dispersive materials which behave as metamaterials (with negative index of refraction) in some range of frequencies. Their constitutive laws (links between electric and magnetic fields and the related inductions) are described in terms of ordinary differential equations (harmonic oscillators). We have studied Maxwell’s equations in such materials and in particular proposed an energy preserving space-time discretization scheme based on an extension of classical methods (mixed finite elements in space, leap-frog type schemes in time). This has been implemented in a code including PMLs for the treatment of unbounded domains. Various numerical experiments have been performed. They illustrate the spectacular dispersive effects of Lorentz materials and allow us to recover the expected focalisation phenomena by an interface between a metamaterial and a standard dielectric one. This code will be an essential tool for the further investigation of more theoretical (and challenging) issues such as the limiting amplitude principle in this context.

6.1.8. Evolution problems in perturbed infinite periodic media

**Participants:** Julien Coatléven, Sonia Fliss.

First, as a part of the PhD of J. Coatléven, based on the former method to solve linear evolution problems in locally perturbed infinite periodic strips through the construction of transparent boundary conditions, a method has been developed to solve its natural geometric extension, i.e. the case of a locally perturbed line defect in an infinite periodic media. The method is again based on a semi-discretization in time of the problem on the whole infinite periodic media, and a generalization of the Lippmann-Schwinger equation approach we have developed for the treatment of this kind of geometries but for harmonic wave problems. At each time step, the solution is written as the sum of the solution of a time step in the unperturbed line defect and a contribution of the perturbation due not only to the current time step but also to all time steps involved in the numerical scheme. This intricate of time steps requires a careful treatment of the Lippmann-Schwinger equation, and in particular of the source term. Using the Floquet-Bloch transform, the computation of all the quantities involved can be reduced to the resolution of a few time steps of linear evolution problems in locally perturbed infinite periodic strips, where we can use the former method. As in the harmonic case, the discretization of the inverse Floquet-Bloch transform is done using appropriate quadrature rules, whereas the space discretization requires classical finite elements. The theoretical basis as well as the numerical analysis of this method are well understood, and the method has been successfully tested numerically. For instance, it can be shown, and checked numerically, that in the particular case of the wave equation, if one uses enough quadrature points (depending on the length of the time interval), then this quadrature creates no error of approximation.

For parabolic problems set in locally perturbed periodic media, we have developed another approach to determine the time-domain DtN operator. The principle is to apply the Laplace Transform in time to the equation and use the construction of the DtN operator for stationary equations. The main difficulty is the computation of the inverse of the Laplace Transform, more precisely to understand how to deal with the unbounded interval of integration and the choice of the discretization of the laplace variable. To deal with the first difficulty for waveguide problem, we have studied the asymptotic behavior of the DtN operator in the laplace domain when the laplace variable tends to $p_0 \pm \infty$. To deal with the second difficulty, we have used the Z-Transformation and its properties. The numerical study is still in progress. This work enters in the framework of the ANR PRoject MicroWave (Sonia Fliss is an external collaborator), in collaboration with Karim Ramdani (Institut Elie Cartan de Nancy, UMR CNRS 7502), Christophe Besse and Ingrid Violet (Laboratoire Paul Painlevé, UMR CNRS 8524).

6.1.9. Modeling and numerical simulation of a piano.

**Participants:** Juliette Chabassier, Patrick Joly.
This work is developed in collaboration with Antoine Chaigne (UME, ENSTA). The purpose of this work is the time domain modeling and numerical simulation of a piano. We aim at explaining the vibratory and acoustical behavior of the piano, by taking into account the main elements that contribute to sound production. The soundboard is modeled as a bidimensional thick, orthotropic, heterogeneous, frequency dependent damped plate, using Reissner Mindlin equations. The vibroacoustics equations allow the soundboard to radiate into the surrounding air, in which we wish to compute the complete acoustical field around the perfectly rigid rim. The soundboard is also coupled to the strings at the bridge, where they form a slight angle from horizontal. Each string is modeled by a one dimensional damped system of equations, taking into account not only the transversal waves excited by the hammer, but also the stiffness thanks to shear waves, as well as the longitudinal waves arising from geometric nonlinearities. The hammer is given an initial velocity that projects it towards a choir of strings, before being repelled. The interacting force is a nonlinear function of the hammer compression.

The final piano model that is discretized is a coupled system of partial differential equations, each of them exhibiting specific difficulties (nonlinear nature of the string system of equations, frequency dependent damping of the soundboard, great number of unknowns required for the acoustic propagation), in addition to couplings’ inherent difficulties. On the one hand, numerical stability of the discrete scheme can be compromised by nonlinear and coupling terms. A very efficient way to guarantee this stability is to construct a numerical scheme which ensures the conservation (or dissipation) of a discrete equivalent of the continuous energy, across time steps. A major contribution of this work has been to develop energy preserving schemes for a class of nonlinear systems of equations, in which enters the string model. On the other hand, numerical efficiency and computation time reduction require that the unknowns of each problem’s part, for which time discretization is specific, hence different, be updated separately. To achieve this artificial decoupling, adapted Schur complements are performed after Lagrange multipliers are introduced.

The potential of this time domain piano modeling is emphasized by realistic numerical simulations. Beyond greatly replicating the measurements, the program allows us to investigate the influence of physical phenomena (string stiffness or nonlinearity), geometry or materials on the general vibratory behavior of the piano, sound included. Spectral enrichment, << phantom partials >> and nonlinear precursors are clearly revealed when large playing amplitudes are involved, highlighting how this approach can help better understand how a piano works.

The main contributions of this year have been the following:

- to write the acoustic propagation as a first order system of equation, involving the physical sound pressure (as opposed to before, when its primitive had to be considered) and the acoustical velocity: this allowed us to artificially bound the computation domain with Perfectly Matched Layers.
- to decrease the numerical computation times via a massive parallelization of the code: parallel modal search for the soundboard, parallel dense matrix-vector product for the vibroacoustic equations in the modal basis, parallel resolution of the 3D acoustic propagation, multi-threaded computation of the strings’ problem.
- to perform realistic computations, which have provided physically relevant numerical simulations (see Figure 3).

Several measurements have also been conducted on a grand piano in order to provide realistic values of parameters and calibrated data to compare simulation with.

6.2. Time-harmonic diffraction problems

6.2.1. Numerical computation of variational integral equation methods

Participants: Marc Lenoir, Nicolas Salles.
The dramatic increase of the efficiency of the variational integral equation methods for the solution of scattering problems must not hide the difficulties remaining for an accurate numerical computation of some influence coefficients, especially when the panels are close and almost parallel.

The formulas have been extended to double layer potentials and, for self influence coefficients, to affine basis functions. Their efficiency for the solution of Maxwell equations has been proved in the framework of a collaboration with CERFACS. The redaction of a paper devoted to the case of parallel panels has been completed and submitted to SIAM J. Sci. Comp. Some preliminary work on the numerical integration of the regular part of the integrand has been undertaken in the context of an internship.

6.2.2. Fast Multipole Method for Viscoelastodynamics

Participants: Marc Bonnet, Stéphanie Chaillat.

This work is done in collaboration with Eva Grasso (LMS, Ecole Polytechnique) and Jean-François Semblat (IFSTTAR). We have extended the single- and multi-domain time-harmonic elastodynamic multi-level fast multipole BEM (Boundary Element Method) formulations to the case of weakly dissipative viscoelastic media [21]. The underlying boundary integral equation and fast multipole formulations are formally identical to that of elastodynamics, except that the wavenumbers are complex-valued due to attenuation. Attention was focused on evaluating the multipole decomposition of the visco-elastodynamic fundamental solution, involving complex-valued wavenumbers. As a result, a damping-dependent modification of the selection rule for the multipole truncation parameter was proposed and assessed on 3D single-region and multi-region visco-elastodynamic examples involving up to about $3 \times 10^5$ boundary nodal unknowns.

6.2.3. Formulation and Fast Evaluation of the Multipole Expansions of the Elastic Half-Space Fundamental Solutions

Participants: Marc Bonnet, Stéphanie Chaillat.
This ongoing work is concerned with a formulation and computation algorithm for the elastodynamic Green’s tensor for the traction-free half-space allowing its use within a Fast Multipole Boundary Element Method (FM-BEM). Due to the implicit satisfaction of the traction-free boundary condition achieved by the Green’s tensor, discretization of (parts of) the free surface is no longer required. Unlike the full-space fundamental solution, the elastodynamic half-space Green’s tensor cannot be expressed in terms of usual kernels such as $e^{ikr}/r$ or $1/r$. Its multipole expansion thus cannot be deduced from known expansions, and is formulated in this work using a spatial two-dimensional Fourier transform approach. The latter achieves the separation of variables which is required by the FMM. A key numerical issue, upon which current work is focused, is concerned with the definition of an efficient numerical quadrature for the evaluation of the inverse Fourier transform, whose integrand is both singular and oscillatory, as classical Gaussian quadratures would perform poorly, fail or require unacceptably large number of quadrature points.

6.2.4. **Multiple scattering by small scatterers**

**Participants:** Maxence Cassier, Christophe Hazard.

We consider the scattering of an acoustic time-harmonic wave by an arbitrary number of sound-soft obstacles located in a homogeneous medium. When the size of the obstacles is small compared with the wavelength, the numerical simulation of such a problem by classical methods (e.g., integral equation techniques or methods based on a Dirichlet-to Neumann map) can become highly time-consuming, particularly when the number of scatterers is large. In this case, the use of an asymptotic model may reduce considerably the numerical cost. Such a model was introduced by Foldy and Lax in the middle of the last century to study multiple isotropic scattering in a medium which contains randomly distributed small scatterers. Their asymptotic model is based on the fact that the scattered wave can be approximated by a wave emitted by point sources placed at the centers of the scatterers; the amplitudes of the sources are calculated by solving a linear system which represents the interactions between the scatterers. Nowadays, the Foldy–Lax model is still used in numerous physical and numerical applications to approximate the scattered wave in a deterministic media. But to the best of our knowledge, there was no mathematical justification of this asymptotic model. We have proposed such a justification which provides local error estimates for the two-dimensional problem in the case of circular obstacles. An article on this subject has been recently submitted to Wave Motion.

6.2.5. **Harmonic wave propagation in locally perturbed infinite periodic media**

**Participants:** Julien Coatléven, Sonia Fliss, Patrick Joly.

A part of the PhD of J. Coatléven consists in developing a method for solving harmonic wave problems with locally perturbed line defects in periodic media. For the treatment of these unbounded defects, which are structured apart from a local perturbation, a new approach has been developed, based on a perturbation principle. The solution is written as the sum of a solution corresponding to the unperturbed line defect and a contribution of the local perturbation. This decomposition leads to a generalization of the so-called Lippmann-Schwinger equation, whose coefficients are computed through their Floquet-Bloch transform, which leads to solve wave-guide problems, these last problems being solved using the transparent boundary condition method developed during S. Fliss’s PhD. The discretization of the inverse Floquet-Bloch transform is done using appropriate quadrature rules, whereas the space discretization requires classical finite elements. The theoretical basis as well as the numerical analysis of this method are well understood, and the method has been successfully tested numerically. In particular, the theoretical convergence estimates have been checked in practice, and the method has a satisfying behavior in limit cases not fully covered by the theory, such as the non-absorbing case.

Concerning the non absorbing case, the question of the limiting absorption principle has been treated for locally perturbed periodic media with particular assumptions. In this case, we are studying the behavior of the solution at large distance of the local perturbation.

6.2.6. **Time harmonic aeroacoustics**

**Participants:** Anne-Sophie Bonnet-Ben Dhia, Jean-François Mercier.
We are still working on the numerical simulation of the acoustic scattering and radiation in presence of a mean flow. This is the object of the ANR project AEROSON, in collaboration with Florence Millot and Sébastien Pernet at CERFACS, Nolwenn Balin at EADS and Vincent Pagneux at the Laboratoire d’Acoustique de l’Université du Maine. The main recent improvements concern: the consideration of ducts with treated boundaries and the development of an alternative model to Galbrun’s equation.

**Treated boundaries**

Our aim is to extend the time harmonic equation of Galbrun to take into account acoustically treated boundaries. Such boundaries are generally described by the Myers boundary condition. Since this condition is naturally expressed in terms of Galbrun’s unknown, the displacement $u$, Galbrun’s equation easily extends to treated boundaries. However we face a difficulty: the original equation of Galbrun leads to a non coercive problem. For rigid boundaries, considering an augmented variational formulation leads to well-posedness. But this approach does not work anymore for treated boundaries.

We have improved our understanding of this difficulty. We are now convinced that the Augmented Galbrun’s equation combined with Myers condition leads to an ill-posed problem. More precisely source terms for which the solution of Galbrun’s equation does not belong to standard functional spaces exist. Such source terms are very particular: located on the treated boundary and singular. This is why during the numerical validations performed at Cerfacs, for “standard” source terms ($L^2$ functions compactly supported in the fluid) we did not get any problem.

To get a well-posed problem, Myers boundary condition, which just requires the normal displacement to belong to $L^2$, must be regularized. It can be achieved by requiring the tangential derivative of $u \cdot n$ to belong to $L^2$. We have also understood that less regularity is sufficient to get a well-posed problem, as it is the case if the fluid is in contact of an elastic medium (such interface cannot be described by a Myers condition since it is necessarily non-local). In particular in the case of a uniform flow well-posedness is proved for a sufficiently slow flow.

**Alternative to Galbrun’s model**

We have kept on considering the model of Goldstein’s equations, alternative to Galbrun’s equation. The Goldstein’s equations couples two unknowns: the velocity potential $\varphi$ and a vectorial unknown $\xi$. $\varphi$ satisfies a modified Helmholtz’s equation with variable coefficients linked to the flow, in which $\xi$ is added as a source term. $\xi$ satisfies a transport equation coupled to the velocity potential. This new model facilitate the treatment of 3D problems since Galbrun’s equation requires to introduce many unknowns. Moreover the vectorial unknown in Goldstein’s formulation vanishes in the areas where the flow is potential which is interesting since realistic flows are mainly potential, the non-potential areas being located near the boundaries or behind obstacles.

As it is the case to calculate the vorticity $\psi = \text{curl}u$, used to regularize Galbrun’s equation, a Discontinuous Galerkin (DG) discretization is used to determine $\xi$ and numerical simulations have been performed at Cerfacs. We have also developed an alternative method allowing to solve Goldstein’s equations with simple Lagrange Finite Element and this was the object of Jean-Emmanuel Lauzet’s internship. We have developed a method combining the Streamline Upwind Petrov Galerkin (SUPG) scheme to discretize the Goldstein’s model with the introduction of PML to bound the calculation domain. To test the efficiency of the method, a non-potential flow has been determined analytically. It consists of a lid-cavity flow connected to a uniform flow in a duct. The viscous cavity flow is solution of the Stoke’s equation and is determined in a rectangular domain by a modal method.

**6.2.7. Modeling of meta-materials in electromagnetism**

**Participants:** Anne-Sophie Bonnet-Ben Dhia, Patrick Ciarlet, Lucas Chesnel.

A collaboration with Eric Chung (Chinese Univ. of Hong Kong) and Xavier Claeys (ISAE).
Meta-materials can be seen as particular media whose dielectric and/or magnetic constant are negative, at least for a certain range of frequency. This type of behavior can be obtained, for instance, with particular periodic structures. Of special interest is the transmission of an electromagnetic wave between two media with opposite sign dielectric and/or magnetic constants. As a matter of fact, applied mathematicians have to address challenging issues, both from the theoretical and the discretization points of view.

The first topic we considered a few years ago was: when is the (simplified) scalar model well-posed in the classical $H^1$ framework? It turned out this issue could be solved with the help of the so-called $T$-coercivity framework. While numerically, we proved that the (simplified) scalar model could be solved efficiently by the most "naive" discretization, still using $T$-coercivity.

Recently, we have been able to provide sharp conditions for the $T$-coercivity to hold in general 2D and 3D geometries (with L. Chesnel), which involve explicit estimates in simplified geometries together with localization arguments. We then analyzed the discretization of the scalar problem with a classical, $H^1$ conforming, finite element method, and proved the convergence under the same sharp conditions (with L. Chesnel). We also showed that the problem can be solved with the help of a Discontinuous Galerkin discretization, which allows one to approximate both the field and its gradient (with E. Chung).

Last (with L. Chesnel and X. Claeys), we investigated the case of a 2D corner which can be ill-posed (in the classical $H^1$ framework). Using the Mellin transform, we showed that a radiation condition at the corner has to be imposed to restore well-posedness. Indeed there exists a wave which takes an infinite time to reach the corner: this "black hole" phenomenon is observed in other situations (elastic wedges for example).

As a second topic, we studied the transmission problem in a purely 3D electromagnetic setting from a theoretical point of view: to achieve well-posedness of this problem, we had to proceed in several steps, proving in particular that the space of electric fields is compactly embedded in $L^2$. For that, we had to assume that the interface is "sufficiently smooth", excluding in particular corners. With L. Chesnel, we have been able to remove this assumption, so that we can solve the problem around an interface with corners. It turns out the $T$-coercivity framework can be applied once more, under the same assumptions as for the scalar model. In the process, we recover more compact embedding results.

6.2.8. Numerical MicroLocal Analysis

Participants: Jean-David Benamou, Francis Collino, Simon Marmorat.

Numerical microlocal analysis of harmonic wavefields is based on a family of linear filters using Bessel functions and applied to wave data collected on a circle of fixed radius $r_0$ around the observation point $x_0$ where we want to estimate the Geometric Optics/ High Frequency components. The data can easily be reconstructed from more conventional line array or grid geometry. The output is an angular function presenting picks of amplitudes in the direction angles of rays.

The original NMLA algorithm relied on a local plane wave assumption for the data. For arbitrary waves, it meant linearization errors and accuracy limitations. Also, only the directions of the (multiple) rays are recovered but the traveltime and amplitudes are not reliably computed. We recently introduced a new "impedant" observable which allows to prove a stability theorem. Numerical results confirm that the new NMLA filter is robust to random and correlated noise.

Using asymptotic expansion on NMLA filtered point sources data, we designed a correction method for the angle which also estimates the wavefront curvature. It can be used to correct the linearization errors mentioned above and provides a second order correction in the Taylor approximation of the traveltime.

The parameters of the method (size of observation circle, discretization) are automatically optimized and a posteriori quantitative error on angles and curvature are available. Numerical studies validate the stability result and confirm the superior accuracy of the curvature corrected NMLA version over image processing methods.
When some bandwith is available we can also compute the traveltime. The amplitude remains polluted by phase errors. Its determination is still open.

6.3. Absorbing boundary conditions and absorbing layers

6.3.1. Perfectly matched transmission problem with absorbing layers : application to anisotropic acoustics

Participant: Sébastien Impériale.

This work has been carried out in collaboration with Edouard Demaldent from CEA-LIST. We have worked on an original approach to design perfectly matched layers (PML) for transient wave equations. This approach is based, first, on the introduction of a modified wave equation and, second, on the formulation of general “perfectly matched” transmission conditions for this equation. The stability of the transmission problem is discussed by way of the adaptation of a high frequency stability (necessary) condition, and we apply our approach to define PML suited for the anisotropic wave equation. A variational formulation of the problem is then developed. It includes a Lagrange multiplier at the interface between the physical and the absorbing domains to deal with the “perfectly matched” transmission conditions. We have carried out numerical results in 2D and 3D that first show the validity of our approach in term of stability and accuracy and the efficiency when using constant damping coefficients combined with high order elements. This work has been submitted for publication.

6.3.2. On the stability in PML corner domains

Participant: Eliane Bécache.

In collaboration with Andres Prieto, from the university of Santiago de Compostella.

We have finalized our work on the stability of the discretization of PMLs in the corners and submitted a paper (see preprint [24]).

6.3.3. High-order Absorbing Boundary Conditions for anisotropic elastodynamics

Participants: Daniel Baffet, Eliane Bécache.

This work is done in collaboration with Daniel Baffet, PhD student of Dan Givoli, at the Technion University in Haifa (Israel) and has started during a visit of Daniel at Poems.

The aim is to design new efficient and stable absorbing boundary conditions for anistropic materials. It is known that the anisotropy introduces a specific difficulty, for ABC as well as for PMLs, in particular for models which involve inverse modes, i.e. waves for which the phase velocity and the group velocity propagate in opposite directions (with respect to the boundary). This has given rise to specific treatment for scalar models but for anisotropic elastodynamics, there are some materials for which no satisfactory solution exist. for these materials however, we can design a low order boundary condition, which is proved to be stable via an energy estimate. We have started to investigate several ways to design higher-order boundary conditions. The main difficulty is to show whether these boundary conditions are stable or not...

6.3.4. Dirichlet to Neumann map with overlap for waveguides

Participants: Anne-Sophie Bonnet-Ben Dhia, Sonia Fliss, Geoffrey Martinache, Antoine Tonnoir.

For scattering problems in acoustic waveguides, a usual approach consists in restricting the computation to a bounded domain containing the sources and the perturbations, using transparent boundary conditions on the artificial boundaries. These conditions are written by using the so-called Dirichlet-to-Neumann maps which can be expressed thanks to a modal decomposition.
An iterative solution of the related linear system can be seen as a domain decomposition formulation without overlapping, where one domain is the bounded region and the other one is infinite. This iterative method does not converge necessarily. A classical idea is to consider a domain decomposition method with overlapping. In this work, we find the equivalent of this method in terms of a new Dirichlet-to-Neumann operator which links the trace of the solution on a section of the waveguide to the normal trace on a different one. This operator can also be expressed analytically via a modal decomposition. Its main advantage is that, because of the overlapping, it becomes compact and this is exactly why we think an iterative resolution has more chance to converge. Other advantages will appear with the elasticity application. Indeed, in the formulation of the transparent boundary condition without overlapping, appears a lagrange multiplier which makes the resolution more costly. This additional unknown will be avoided with an overlap.

For now, the theory is done for the scalar acoustic waveguide and the method has been implemented in the Melina code. The extension to the elastic case is in progress.

6.3.5. An alternative to DTN maps in waveguides

Participants: Anne-Sophie Bonnet-Ben Dhia, Guillaume Legendre.

We are interested by the treatment of the radiation condition at infinity for the numerical solution of a problem set in an unbounded waveguide. We have proposed an alternative to the classical approach involving a modal expression of Dirichlet-to-Neumann (DtN) operators. This new method is particularly simple to implement since it only requires to solve several times a boundary value problem with local boundary conditions. In the case of a an acoustic waveguide, we have proved that the corresponding approximate solution is comparable in accuracy to the one obtained by truncating the infinite series in the DtN maps. The number of linear systems to invert, which has to be greater than the number of propagative guided modes, can be significantly reduced by combining the approach with the perfectly matched layer (PML) technique. It works even in elastic waveguides, despite existence of the so-called backward waves which are known to make the PMLs inefficient, when used alone.

6.4. Waveguides, resonances, and scattering theory

6.4.1. Modelling of non-homogeneous lossy coaxial cable for time domain simulation.

Participants: Sébastien Impériale, Patrick Joly.

In this work, we focus on the time-domain simulation of the propagation of electromagnetic waves in non-homogeneous lossy coaxial cables. This question has been motivated by our collaboration with CEA-LIST about the numerical modeling of non-destructive testing (for the detection of cracks in metallic bodies for instance) by ultra-sounds and more precisely the modeling of piezo-electric transducers (see section 6.1.6). The complete description of such devices often requires an accurate modeling of the supply process, which includes the propagation of the electric current along coaxial cables. This question appears as an independent sub-problem.

The main characteristic of such coaxial cables is that their transverse directions are very small with respect to their length as well as the wavelength. As a consequence, one would like to use a simplified 1D model as an effective (or homogenized) model for electromagnetic propagation. In this work, we construct and justify rigorously such a model by way of an asymptotic analysis of time harmonic 3D Maxwell’ s equations in such a structure. The effective model appears as a generalized wave equation with additional time convolution terms that take into account electric and magnetic losses. By this way, we justify and extend some models proposed in the electrical engineering literature, in particular the well-known telegraphist’s equation. The properties of our limit model in time domain has been analyzed and a stable discretization process has been proposed. Numerical simulation in academic simulations exhibit some exotic phenomena of "dispersive dissipation".

Our further investigations of the subject will be developed in the framework of the new ANR project SODDA in collaboration with the team Sysiphe (M. Sorine).
6.4.2. Study of lineic defect in periodic media

**Participant:** Sonia Fliss.

This work deals with one dimensional infinite perturbation - namely line defects - in periodic media. In optics, such defects are created to construct an (open) waveguide that concentrates light. The existence and the computation of the eigenmodes is a crucial issue. This is related to a selfadjoint eigenvalue problem associated to a PDE in an unbounded domain (in the directions orthogonal to the line defect), which makes both the analysis and the computations more complex. Using a Dirichlet-to-Neumann (DtN) approach, we show that this problem is equivalent to one set on a small neighborhood of the defect. In opposition to existing methods, this method is exact but there is a price to be paid: the reduction of the problem leads to a nonlinear eigenvalue problem of a fixed point nature. An article presenting the method and its properties is being written, the numerical study is in progress in collaboration with Kersten Schmidt and Dirk Klindworth from the Technische Universität Berlin.

Our further investigations of the subject will be developed in the framework of a new DGA project.

6.4.3. A new approach for the numerical computation of non linear modes of vibrating systems

**Participants:** Anne-Sophie Bonnet-Ben Dhia, Jean-François Mercier.

A collaboration with Cyril Touzé and François Blanc (Unité de Mécanique, ENSTA). The simulation of vibrations of large amplitude of thin plates or shells requires the expensive solution of a non-linear finite element model. The main objective of the proposed study is to develop a reliable numerical method which reduces drastically the number of degrees of freedom. The main idea is the use of the so-called non-linear modes to project the dynamics on invariant subspaces, in order to generate accurate reduced-order models. Cyril Touzé from the Unité de Mécanique of ENSTA has derived an asymptotic method of calculation of the non-linear modes for both conservative and damped systems. But the asymptotically computed solution remains accurate only for moderate amplitudes. This motivates the present study which consists in developing a numerical method for the computation of the non-linear modes, without any asymptotic assumption. This is the object of a collaboration with Cyril Touzé, and new results have been obtained during the post-doc of François Blanc in the Unité de Mécanique of ENSTA. The partial differential equations defining the invariant manifold of the non-linear mode are seen as a vectorial transport problem: the variables are the amplitude and the phase ($a, \varphi$) where the phase $\varphi$ plays the role of the time. In the case of conservative systems, a finite difference scheme is used and an iterative algorithm is written, to take into account the $2\pi$-periodicity in $\varphi$ which is seen as a constraint. An adjoint state approach has been introduced to evaluate the gradient of the cost function. The method has been validated in a simple example with two degrees of freedom. Good agreement with an alternative method, the continuation of periodic solutions method, has been found. Currently the method is extended to the case of damped systems. The main difficulty is that, due to a change of variables, the $2\pi$-periodicity does not hold anymore and new constraints more complicated to implement must be considered. Numerical implementation is still under progress.

6.5. Asymptotic methods and approximate models

6.5.1. Effective boundary conditions for thin periodic coatings

**Participants:** Mathieu Chamaillard, Patrick Joly.

This topic is developed in collaboration with H. Haddar (DEFI, INRIA Saclay) can be seen as a continuation of the PhD thesis of B. Delourme (see the activity report of last year) on effective transmission conditions for thin rough interfaces. On this last subject, the mathematical analysis of such transmission conditions in the (difficult) case of 3D Maxwell’s equations has been completed and submitted for publication.
We are now coming back to the more traditional issue of effective or approximate boundary conditions for simulating thin periodic coatings at the surface of a diffractive obstacle. This subject has already been more widely investigated, in particular in France (see the works by Achdou, Ammari and their collaborators for instance). However, we attack, with the PhD thesis of M. Chamaillard, supported by a DGA/INRIA scholarship, various new aspects of the problem, namely:

- the treatment of surfaces of general geometry,
- the use of non standard materials, such as ferromagnetic materials, for the coating,
- the construction of higher order impedance conditions.

This is motivated by various recent progress in the domain of stealth technology and we hope to develop a collaboration in this domain with CEA-DAM (CESTA and Le Ripault).

6.5.2. Elastic wave propagation in strongly heterogeneous media

Participants: Patrick Joly, Simon Marmorat.

This subject enters our long term collaboration with CEA-LIST on the development on numerical methods for time-domain non destructive testing experiments using ultra-sounds. This is also the subject of the PhD thesis of Simon Marmorat. Our objective is to develop an efficient numerical approach for the propagation of elastic waves in a medium which is made of many small inclusions / heterogeneities embedded in a smooth (or piecewise smooth) background medium, without any particular assumption (such as periodicity) on the spatial distribution of these heterogeneities. Our idea is to exploit the smallness of the inclusions (with respect to the wavelength in the background medium) to derive a simplified approximate model in which each inclusion would be described by very few parameters (functions of time) coupled to the displacement field in background medium for which we could use a computational mesh that ignores the presence of the heterogeneities. For deriving such a model, we intend to use and adapt the asymptotic methods previously developed at Poems (such as matched asymptotic expansions).

6.5.3. Approximate models in aeroacoustics

Participants: Anne-Sophie Bonnet-Ben Dhia, Patrick Joly, Guillaume Legendre, Ricardo Weder.

This topic concerns the 2D acoustic propagation in presence of a mean flow, modeled for instance by the Galbrun equation. We had previously derived effective boundary conditions taking into account the boundary layers of the mean flow near a rigid or treated boundary. These boundary conditions are in general non local with respect to the normal coordinate inside the boundary layer. However when the Mach profile in the shear layer is piecewise linear, the condition can be replaced by a system of 1D advection equations, which are coupled with the Galbrun equation in the 2D domain. We have derived a variational formulation for this model, in time-harmonic regime and for the case of a rigid boundary. This formulation has been implemented in the Melina code: the first results are promising and the validation, by comparison to the solution obtained by a full discretization of the shear layer, is in progress.

6.6. Imaging and inverse problems

6.6.1. Sampling methods in waveguides

Participants: Laurent Bourgeois, Eric Lunéville, Alexandre Routier.

We have derived a modal formulation of sampling methods (both the Linear Sampling Method and the Factorization Method) in an acoustic waveguide when the obstacle to recover is a set of cracks. This was the subject of the Master internship of Alexandre Routier. For such particular obstacle, we have analyzed the importance of the test function we introduce in the sampling method if we a priori know the type of boundary condition we have on the lips of the crack, both from the theoretical and the numerical point of view. Besides we have proved that in our modal formulation, the Factorization Method is applicable by using the same data as those used in the Linear Sampling Method, which is a novelty as concerns sampling methods in waveguides. The Linear Sampling Method has been extended to the elastic case, for which the usual obstacle is a set of traction free cracks. This makes the choice of the test function crucial, and we have emphasized such fact on numerical examples.
6.6.2. **Inverse scattering with generalized impedance boundary conditions**

**Participants:** Laurent Bourgeois, Nicolas Chaulet.

This work is a collaboration between POEMS and DEFI projects (more precisely Housssem Haddar) and constitutes the subject of the PhD thesis of N. Chaulet. We are concerned with the identification of some obstacle and some Generalized Impedance Boundary Conditions (GIBC) on the boundary of such obstacle from far field measurements. The GIBCs are approximate models for thin coatings or corrugated surfaces. During this last year, we have completed the computation of the partial derivatives of the far field with respect to the unknowns, among which is the boundary of the obstacle, and we have implemented many numerical experiments. In particular, we have shown the efficiency of the method consisting in approximating a perfect conductor which is coated with a thin dielectric layer of variable width by a second order GIBC in order to retrieve the obstacle, as well as the refraction index and the width of the layer.

6.6.3. **Detection of targets using time-reversal**

**Participants:** Maxence Cassier, Patrick Joly, Christophe Hazard.

This topic concerns the studies started last year about time-reversal in the context of Maxence Cassier’s thesis. The main question is to generate a time-dependent wave that focuses on one given scatterer not only in space, but also in time. Our recent works concern two items. On one hand, we have proposed a way to construct such a focusing wave which does not require an a priori knowledge of the location of the obstacle. This wave is represented by a suitable superposition of the eigenvectors of the so-called time-reversal operator in the frequency domain. Numerical results show the focusing properties of such a wave. On the other hand, we try to understand how to translate the physical idea of “focusing” into mathematical terms. We have proposed an energy criterion which can be used in numerical experiments in order to evaluate the quality of the focus. The question is to relate such a criterion with the construction of the above mentioned focusing wave. Works on this topic are in progress.

6.6.4. **Interior transmission problem**

**Participants:** Anne-Sophie Bonnet-Ben Dhia, Lucas Chesnel.

This work is a collaboration with Housssem Haddar from the DEFI project. The interior transmission problem plays an important role in the inverse scattering theory for inhomogeneous media. In particular, it arises when one is interested in the reconstruction of an inclusion embedded in a background medium from multi-static measurements of diffracted fields at a given frequency. Physically, it is important to prove that, for a given frequency, there are no waves which do not scatter. Mathematically, this last property boils down to say that the frequency is not a transmission eigenvalue, that is, an eigenvalue of the interior transmission problem. An important issue is to prove that transmission eigenvalues form at most a discrete set with infinity as the only accumulation point. This is not trivial because the operator associated with this problem exhibits a sign changing in its principal part and its study is not standard. Using the T-coercivity approach, we proved the discreteness under relatively weak assumptions. In particular, the simple technique we proposed allows to treat cases, which were not covered by existing methods, where the difference between the inclusion index and the background index changes sign.

6.6.5. **Flaw identification using elastodynamic topological derivative**

**Participant:** Marc Bonnet.

In collaboration with Cédric Bellis (Columbia Univ. USA), Bojan Guzina (Univ. of Minnesota, USA). The concept of topological derivative (TD) quantifies the perturbation induced to a given cost functional by the nucleation of an infinitesimal flaw in a reference defect-free body, and may serve as a flaw indicator function. In this work, the TD is derived for three-dimensional crack identification exploiting over-determined transient elastodynamic boundary data. This entails in particular the derivation of the relevant polarization tensor, here given for infinitesimal trial cracks in homogeneous or bi-material elastic bodies. Simple and efficient adjoint-state based formulations are used for computational efficiency, allowing to compute the TD field for arbitrarily shaped elastic solids. The latter is then used as an indicator function for the spatial location of the sought
crack(s). This approach, which allows a qualitative reconstruction of cracks in terms of their location but also their orientation (utilizing the fact that the polarization tensor depends on the normal to the trial small crack), has been implemented within a conventional FEM platform. A standard Newmark unconditionally-stable time-marching scheme is used for simulating data, and for computing the free and adjoint solutions used in the evaluation of the TD field. Extensive 3D time-domain numerical experiments for the detection of cracks buried either in a homogeneous pipe-like structure or on the interface between two sandwiched plates highlight its usefulness and performance. The application of TD to flaw identification has thus far rested upon a heuristic basis. Its justification in limiting situations such as the Born approximation is currently being investigated.

6.6.6. Topological derivative in anisotropic elasticity

Participant: Marc Bonnet.

In collaboration with Gabriel Delgado (CMAP, Ecole Polytechnique). This work addresses the current lack of a comprehensive treatment of the topological derivative for anisotropic elasticity, by addressing the case where both the background material and the trial small inhomogeneity have arbitrary anisotropic elastic properties. Accordingly, a formula for the topological derivative of any cost functional defined in terms of regular volume or surface densities depending on the displacement is established, by combining small-inhomogeneity asymptotics and the adjoint solution approach. The latter feature makes the proposed result simple to implement and computationally efficient. Both three-dimensional and plane-strain settings are covered; they differ mostly on details pertaining to the elastic moment tensor. This result achieves a direct generalization to the fully anisotropic case of previously-known formulations for isotropic elasticity. Moreover, the main properties of the EMT, a critical feature of any elastic topological derivative formula, are studied for the fully anisotropic case, generalizing available results on the isotropic case. Finally, further generality is achieved by also deriving the topological derivative of strain energy-based cost functionals, which depend on the displacement gradient. This case, seldom addressed so far, requires a specific, and separate, treatment. Applications of these results include topology optimization of composite structures (a topic currently pursued by G. Delgado) or flaw identification using experimental data from nondestructive testing.

6.6.7. Energy functionals for elastic medium reconstruction using transient data

Participant: Marc Bonnet.

In collaboration with Wilkins Aquino (Cornell Univ., USA).

Energy-based misfit cost functionals, known in mechanics as error in constitutive relation (ECR) functionals, are known since a long time to be well suited to (electrostatic, elastic,...) medium reconstruction. In this ongoing work, a transient elastodynamic version of this methodology is developed, with emphasis on its applicability to large time-domain finite element modeling of the forward problem. The formulation involves coupled transient forward and adjoint solutions, a fact which greatly hinders large-scale computations. A computational approach combining an iterative treatment of the coupled problem and the adjoint to the discrete Newmark time-stepping scheme is found to perform well on large FE models, making the time-domain ECR functional a worthwhile tool for medium identification.

6.6.8. Accelerated boundary element method for diffuse optical imaging

Participant: Marc Bonnet.

In collaboration with Simon Arridge and Josias Elisee (University College London, UK). Numerical methods for calculating forward models of light propagation in tissue are extensively used in diffuse optical imaging (DOI). DOI involves a Helmholtz-type PDE with a complex-valued wavenumber. It requires modelling large optical regions whose parameters are known and piecewise constant. The boundary element method (BEM) answers this need and avoids the detailed interior meshing of these regions. The single-level Fast Multipole Method has been applied for solving the DOI governing equation, allowing substantial reduction of computational costs. The enhanced practicability of the BEM in DOI was demonstrated through test examples on single-layer problems, where two-digit reduction factors on solution time are achieved, and on a high-resolution version of a three-layered neonate’s head.
6.6.9. High-Velocity Estimates and Inverse Scattering for Quantum N-Body Systems with Stark Effect

Participants: Ricardo Weder, Gerardo Daniel Valencia.

In an $N$–body quantum system with a constant electric field, by inverse scattering, we uniquely reconstruct pair potentials, belonging to the optimal class of short-range potentials and long-range potentials, from the high-velocity limit of the Dollard scattering operator. We give a reconstruction formula with an error term.

6.6.10. Small-Energy Analysis for the Matrix Schrödinger Operator on the Half-Line

Participant: Ricardo Weder.

In collaboration with Tuncay Aktosun (University of Texas Arlington) and Martin Klaus (Virginia Tech). The matrix Schrödinger equation with a selfadjoint matrix potential is considered on the half line with the most general selfadjoint boundary condition at the origin. When the matrix potential is integrable and has a first moment, it is shown that the corresponding scattering matrix is continuous at zero energy. An explicit formula is provided for the scattering matrix at zero energy. The small-energy asymptotics are established also for the related Jost matrix, its inverse, and various other quantities relevant to the corresponding direct and inverse scattering problems.

6.7. Other topics

6.7.1. Fast non-overlapping Schwarz domain decomposition methods for the neutron diffusion equation

Participant: Patrick Ciarlet.

A collaboration with Erell Jamelot (CEA Saclay/DEN). Investigating numerically the steady state of a nuclear core reactor can be very expensive, in terms of memory storage and computational time. In order to address both requirements, one can use a domain decomposition method, which is then implemented on a parallel computer.

We model the problem using a mixed approach, which involves a scalar flux and a vector current. The equivalent variational formulation is then discretized with the help of Raviart-Thomas-Nédélec finite elements.

The domain decomposition method is based on the Schwarz iterative algorithm with Robin interface conditions to handle communications. This method is analyzed from the continuous to the discrete point of views: well-posedness, convergence of the finite element method, optimality of the parameter appearing in the Robin interface condition and algorithms. Numerical experiments carried out on realistic 3D configurations using the APOLLO3©code (of CEA/DEN) show the parallel efficiency of the algorithm.

6.7.2. Equivalent local boundary conditions for the Monge Kantorovitch Mass Transport problem

Participant: Jean-David Benamou.

This work is done in collaboration with Adam Oberman Brittany Froese from Simon Fraser University, Vancouver. In the last 20 years, the Monge Kantorovich Optimal Transport problem (OTP) and its relationship with Partial Differential Equations (PDE) experienced a spectacular research revival (the Fields medal was awarded to Cedric Villani in 2010 partly for his contributions to OTP). Applications appeared in fields as diverse as meteorology, medical image processing, astronomy and economy. This new area offers numerical challenges which go beyond current knowledge in the field. Novel computational tools are needed. The OTP can also be reformulated as a Monge-Ampere (MA) PDE with non standard/non local boundary conditions (BC). We would like to use the new and efficient wide-stencil Finite Difference MA solver developed by Oberman and Froese but we do not know how to deal with the OTP equivalent BCs. We pursue the investigation of the two following innovative strategies : 1. Iteratively construct Neumann BCs such that the solution of the MA equation satisfies the OTP BC in the limit. 2. Merge the data in all of space and design simplified asymptotic local BCs at infinity which can be used to formulate local transparent BC on a truncated domain.
6.7.3. **Topological Effects in Quantum Mechanics and High-Velocity Estimates**  
**Participant:** Ricardo Weder.

This work is done in collaboration with Miguel Ballesteros. High-velocity -or high-energy- estimates for scattering solutions to the Schrödinger equation are important for many reasons. For example, in topological effects in quantum mechanics, where the space accessible to the particles has a non-trivial topology, like, for example, in the celebrated magnetic Aharonov-Bohm effect, where an electron is constrained to be on the exterior of a torus that contains a magnetic flux inside. Here, the solution acquires a phase if the electron travels inside the hole of the magnet and, on the contrary, it acquires no phase if the particle travels outside the hole. We obtain precise high-velocity estimates for the scattering solutions, that prove that quantum mechanics actually predicts the existence of the magnetic Aharonov-Bohm effect, under the conditions of the celebrated Tonomura et al. experiments. Moreover, in the case of the electric Aharonov-Bohm effect, we provide precise conditions for the validity of the Aharonov-Bohm Ansatz and we give a rigorous proof that quantum mechanics predicts the existence of this effect.

6.7.4. **Entanglement Creation in Low-Energy Scattering**  
**Participant:** Ricardo Weder.

We study the entanglement creation in the low-energy scattering of two particles of mass, $m_1, m_2$ in three dimensions. We consider a general class of interaction potentials that are not required to be spherically symmetric. The incoming asymptotic state, before the collision, is a product of two normalized Gaussian states with the same variance, $\sigma$, and opposite mean momentum. After the scattering the particles are in the outgoing asymptotic state that is not a product state. We take as a measure of the entanglement created by the collision the purity of one of the particles in the outgoing asymptotic state. In the incoming asymptotic state the purity is one. We provide a rigorous explicit computation, with error bound, of the leading order of the purity at low-energy. The leading order depends strongly in the difference of the masses. The entanglement takes its minimum when the masses are equal, and it increases rapidly with the difference of the masses. It is quite remarkable that the anisotropy of the potential plays no role, on spite of the fact that entanglement is a second order effect.

6.7.5. **Open Scattering Channels in Manifolds with Horns**  
**Participant:** Ricardo Weder.

This work is done in collaboration with Olaf Post (Humboldt University, Berlin) and Rainer Hempel (Mount Allison University, Sackville New Brunswick). In the framework of time-dependent geometric scattering theory, we study the existence and completeness of the wave operators for perturbations of the Riemannian metric for the Laplacian on a complete manifold of dimension $n$. The smallness condition for the perturbation is expressed in purely geometric terms using the harmonic radius; therefore, the size of the perturbation can be controlled in terms of local bounds on the radius of injectivity and the Ricci-curvature, and no global assumption is needed. As an application of these ideas we obtain a stability result for the scattering matrix. As a consequence we find that a scattering channel which interacts with other channels preserves this property under small perturbations.
6. New Results

6.1. Local Component Analysis

Participants: Francis Bach, Nicolas Le Roux.

Kernel density estimation, a.k.a. Parzen windows, is a popular density estimation method, which can be used for outlier detection or clustering. With multivariate data, its performance is heavily reliant on the metric used within the kernel. Most earlier work has focused on learning only the bandwidth of the kernel (i.e., a scalar multiplicative factor). In this paper, we propose to learn a full Euclidean metric through an expectation-minimization (EM) procedure, which can be seen as an unsupervised counterpart to neighbourhood component analysis (NCA). In order to avoid overfitting with a fully nonparametric density estimator in high dimensions, we also consider a semi-parametric Gaussian-Parzen density model, where some of the variables are modelled through a jointly Gaussian density, while others are modelled through Parzen windows. For these two models, EM leads to simple closed-form updates based on matrix inversions and eigenvalue decompositions. We show empirically that our method leads to density estimators with higher test-likelihoods than natural competing methods, and that the metrics may be used within most unsupervised learning techniques that rely on such metrics, such as spectral clustering or manifold learning methods. Finally, we present a stochastic approximation scheme which allows for the use of this method in a large-scale setting [30].

6.2. Weakly Supervised Learning of Foreground-Background Segmentation using Masked RBMs

Participant: Nicolas Le Roux.

Collaboration with: Nicolas Heess (School of Informatics, University of Edinburgh) and John Winn (Machine Learning and Perception, Microsoft research Cambridge).

We propose an extension of the Restricted Boltzmann Machine (RBM) that allows the joint shape and appearance of foreground objects in cluttered images to be modeled independently of the background. We present a learning scheme that learns this representation directly from cluttered images with only very weak supervision. The model generates plausible samples and performs foreground-background segmentation. We demonstrate that representing foreground objects independently of the background can be beneficial in recognition tasks [15].

6.3. Clusterpath: an algorithm for clustering using convex fusion penalties

Participants: Toby Hocking, Francis Bach, Armand Joulin.


We present a new clustering algorithm by proposing a convex relaxation of hierarchical clustering, which results in a family of objective functions with a natural geometric interpretation. We give efficient algorithms for calculating the continuous regularization path of solutions, and discuss relative advantages of the parameters. Our method experimentally gives state-of-the-art results similar to spectral clustering for non-convex clusters, and has the added benefit of learning a tree structure from the data [16].

6.4. Convex and Network Flow Optimization for Structured Sparsity

Participants: Rodolphe Jenatton, Guillaume Obozinski, Francis Bach.

Collaboration with: Julien Mairal (Department of Statistics, University of California, Berkeley).
In [5], we consider a class of learning problems regularized by a structured sparsity-inducing norm defined as the sum of $\ell_2$- or $\ell_\infty$-norms over groups of variables. Whereas much effort has been put in developing fast optimization techniques when the groups are disjoint or embedded in a hierarchy, we address here the case of general overlapping groups. To this end, we present two different strategies: On the one hand, we show that the proximal operator associated with a sum of $\ell_\infty$-norms can be computed exactly in polynomial time by solving a quadratic min-cost flow problem, allowing the use of accelerated proximal gradient methods. On the other hand, we use proximal splitting techniques, and address an equivalent formulation with non-overlapping groups, but in higher dimension and with additional constraints. We propose efficient and scalable algorithms exploiting these two strategies, which are significantly faster than alternative approaches. We illustrate these methods with several problems such as CUR matrix factorization, multi-task learning of tree-structured dictionaries, background subtraction in video sequences, image denoising with wavelets, and topographic dictionary learning of natural image patches.

### 6.5. Multi-scale Mining of fMRI data with Hierarchical Structured Sparsity

**Participants:** Rodolphe Jenatton, Guillaume Obozinski, Francis Bach.

Collaboration with: Alexandre Gramfort, Vincent Michel, Evelyn Eger and Bertrand Thirion (Laboratoire de Neuroimagerie Assistée par Ordinateur (LNAO), CEA: DSV/I2BM/NEUROSPIN, PARIETAL (INRIA Saclay - Ile de France) and Neuroimagerie cognitive, INSERM: U992 – Université Paris Sud – CEA).

Inverse inference, or "brain reading", is a recent paradigm for analyzing functional magnetic resonance imaging (fMRI) data, based on pattern recognition and statistical learning. By predicting some cognitive variables related to brain activation maps, this approach aims at decoding brain activity. Inverse inference takes into account the multivariate information between voxels and is currently the only way to assess how precisely some cognitive information is encoded by the activity of neural populations within the whole brain. However, it relies on a prediction function that is plagued by the curse of dimensionality, since there are far more features than samples, i.e., more voxels than fMRI volumes. To address this problem, different methods have been proposed, such as, among others, univariate feature selection, feature agglomeration and regularization techniques. In this paper, we consider a sparse hierarchical structured regularization. Specifically, the penalization we use is constructed from a tree that is obtained by spatially-constrained agglomerative clustering. This approach encodes the spatial structure of the data at different scales into the regularization, which makes the overall prediction procedure more robust to inter-subject variability. The regularization used induces the selection of spatially coherent predictive brain regions simultaneously at different scales. We test our algorithm on real data acquired to study the mental representation of objects, and we show that the proposed algorithm not only delineates meaningful brain regions but yields as well better prediction accuracy than reference methods [24], [17].

### 6.6. Convergence Rates of Inexact Proximal-Gradient Methods for Convex Optimization

**Participants:** Mark Schmidt, Nicolas Le Roux, Francis Bach.

In [21], we consider the problem of optimizing the sum of a smooth convex function and a non-smooth convex function using proximal-gradient methods, where an error is present in the calculation of the gradient of the smooth term or in the proximity operator with respect to the non-smooth term. We show that both the basic proximal-gradient method and the accelerated proximal-gradient method achieve the same convergence rate as in the error-free case, provided that the errors decrease at appropriate rates. Using these rates, we perform as well as or better than a carefully chosen fixed error level on a set of structured sparsity problems.

### 6.7. Generalized Fast Approximate Energy Minimization via Graph Cuts: Alpha-Expansion Beta-Shrink Moves

**Participant:** Mark Schmidt.
In [20], we present alpha-expansion beta-shrink moves, a simple generalization of the widely-used alpha-beta swap and alpha-expansion algorithms for approximate energy minimization. We show that in a certain sense, these moves dominate both alpha-beta-swap and alpha-expansion moves, but unlike previous generalizations the new moves require no additional assumptions and are still solvable in polynomial-time. We show promising experimental results with the new moves, which we believe could be used in any context where alpha-expansions are currently employed.

6.8. Hybrid Deterministic-Stochastic Methods for Data Fitting

Participant: Mark Schmidt.

Collaboration with: Michael P. Friedlander (University of British Columbia).

Many structured data-fitting applications require the solution of an optimization problem involving a sum over a potentially large number of measurements. Incremental gradient algorithms offer inexpensive iterations by sampling only subsets of the terms in the sum. These methods can make great progress initially, but often slow as they approach a solution. In contrast, full gradient methods achieve steady convergence at the expense of evaluating the full objective and gradient on each iteration. We explore hybrid methods that exhibit the benefits of both approaches. Rate of convergence analysis shows that by controlling the size of the subsets in an incremental gradient algorithm, it is possible to maintain the steady convergence rates of full gradient methods. We detail a practical quasi-Newton implementation based on this approach, and numerical experiments illustrate its potential benefits [29].

6.9. Multi-task regression using minimal penalties

Participants: Matthieu Solnon, Sylvain Arlot, Francis Bach.

In [33] we study the kernel multiple ridge regression framework, which we refer to as multi-task regression, using penalization techniques. The theoretical analysis of this problem shows that the key element appearing for an optimal calibration is the covariance matrix of the noise between the different tasks. We present a new algorithm to estimate this covariance matrix, based on the concept of minimal penalty, which was previously used in the single-task regression framework to estimate the variance of the noise. We show, in a non-asymptotic setting and under mild assumptions on the target function, that this estimator converges towards the covariance matrix. Then plugging this estimator into the corresponding ideal penalty leads to an oracle inequality. We illustrate the behavior of our algorithm on synthetic examples.

6.10. Ask the locals: multi-way local pooling for image recognition

Participants: Nicolas Le Roux, Francis Bach.

Collaboration with: Y-Lan Boureau and Jean Ponce (Willow project-team, INRIA Paris-Rocquencourt) and Yann LeCun (Courant Institute of Mathematical Science (CIMS), New York University).

Invariant representations in object recognition systems are generally obtained by pooling feature vectors over spatially local neighborhoods. But pooling is not local in the feature vector space, so that widely dissimilar features may be pooled together if they are in nearby locations. Recent approaches rely on sophisticated encoding methods and more specialized codebooks (or dictionaries), e.g., learned on subsets of descriptors which are close in feature space, to circumvent this problem. In this work, we argue that a common trait found in much recent work in image recognition or retrieval is that it leverages locality in feature space on top of purely spatial locality. We propose to apply this idea in its simplest form to an object recognition system based on the spatial pyramid framework, to increase the performance of small dictionaries with very little added engineering. State-of-the-art results on several object recognition benchmarks show the promise of this approach [12].
6.11. Trace Lasso: a trace norm regularization for correlated designs

Participants: Edouard Grave, Guillaume Obozinski, Francis Bach.

Using the $\ell_1$-norm to regularize the estimation of the parameter vector of a linear model leads to an unstable estimator when covariates are highly correlated. In this paper, we introduce a new penalty function which takes into account the correlation of the design matrix to stabilize the estimation. This norm, called the trace Lasso, uses the trace norm of the selected covariates, which is a convex surrogate of their rank, as the criterion of model complexity. We analyze the properties of our norm, describe an optimization algorithm based on reweighted least-squares, and illustrate the behavior of this norm on synthetic data, showing that it is more adapted to strong correlations than competing methods such as the elastic net [14].

6.12. Shaping level sets through submodular functions

Participant: Francis Bach.

The concept of parsimony is central in many scientific domains. In the context of statistics, signal processing or machine learning, it may take several forms. Classically, in a variable or feature selection problem, a sparse solution with many zeros is sought so that the model is either more interpretable, cheaper to use, or simply matches available prior knowledge. In this work, we instead consider sparsity-inducing regularization terms that will lead to solutions with many equal values. A classical example is the total variation in one or two dimensions, which leads to piecewise constant solutions and can be applied to various image labelling problems, or change point detection tasks. In this work [9], we follow our earlier approach which consisted in designing sparsity-inducing norms based on non-decreasing submodular functions, as a convex approximation to imposing a specific prior on the supports of the predictors. Here, we show that a similar parallel holds for some other class of submodular functions, namely non-negative set functions which are equal to zero for the full and empty set. Our main instance of such functions are symmetric submodular functions and we show that the Lovász extension may be seen as the convex envelope of a function that depends on level sets (i.e., the set of indices whose corresponding components of the underlying predictor are greater than a given constant). By selecting specific submodular functions, we give a new interpretation to known norms, such as the total variation; we also define new norms, in particular ones that are based on order statistics with application to clustering and outlier detection, and on noisy cuts in graphs with application to change point detection in the presence of outliers.

6.13. Itakura-Saito Nonnegative Matrix Factorization with group sparsity

Participants: Augustin Lefèvre, Francis Bach.

Collaboration with: Cédric Févotte (Laboratoire traitement et communication de l’information (LTCI), CNRS: UMR5141 – Institut Télécom – Télécom ParisTech).

In [18], we propose an unsupervised inference procedure for audio source separation. Components in nonnegative matrix factorization (NMF) are grouped automatically in audio sources via a penalized maximum likelihood approach. The penalty term we introduce favors sparsity at the group level, and is motivated by the assumption that the local amplitude of the sources are independent. Our algorithm extends multiplicative updates for NMF; moreover we propose a test statistic to tune hyperparameters in our model, and illustrate its adequacy on synthetic data. Results on real audio tracks show that our sparsity prior allows to identify audio sources without knowledge on their spectral properties.

6.14. Online algorithms for nonnegative matrix factorization with the Itakura-Saito divergence

Participants: Augustin Lefèvre, Francis Bach.

Collaboration with: Cédric Févotte (Laboratoire traitement et communication de l’information (LTCI), CNRS: UMR5141 – Institut Télécom – Télécom ParisTech).
Nonnegative matrix factorization (NMF) is now a common tool for audio source separation. When learning NMF on large audio databases, one major drawback is that the complexity in time is $O(FKN)$ when updating the dictionary (where $(F;N)$ is the dimension of the input power spectrograms, and $K$ the number of basis spectra), thus forbidding its application on signals longer than an hour. We provide an online algorithm with a complexity of $O(FK)$ in time and memory for updates in the dictionary. We show on audio simulations that the online approach is faster for short audio signals and allows to analyze audio signals of several hours [31].

6.15. A Graph-matching Kernel for Object Categorization

**Participant:** Armand Joulin.

Collaboration with: Olivier Duchenne and Jean Ponce (Willow project-team, INRIA Paris-Rocquencourt).

In [13], we address the problem of category-level image classification. The underlying image model is a graph whose nodes correspond to a dense set of regions, and edges reflect the underlying grid structure of the image and act as springs to guarantee the geometric consistency of nearby regions during matching. A fast approximate algorithm for matching the graphs associated with two images is presented. This algorithm is used to construct a kernel appropriate for SVM-based image classification, and experiments with the Caltech 101, Caltech 256, and Scenes datasets demonstrate performance that matches or exceeds the state of the art for methods using a single type of features.


**Participant:** Guillaume Obozinski.

Collaboration with: Laurent Jacob (Department of Statistics, University of California at Berkeley) and Jean-Philippe Vert (INSERM U900, Mines ParisTech, Institut Curie).

We study in [25] a norm for structured sparsity which leads to sparse linear predictors whose supports are unions of predefined overlapping groups of variables. We call the obtained formulation latent group Lasso, since it is based on applying the usual group Lasso penalty on a set of latent variables. A detailed analysis of the norm and its properties is presented and we characterize conditions under which the set of groups associated with latent variables are correctly identified. We motivate and discuss the delicate choice of weights associated to each group, and illustrate this approach on simulated data and on the problem of breast cancer prognosis from gene expression data.

6.17. Sparse Image Representation with Epitomes

**Participants:** Louise Benoît, Francis Bach.

Collaboration with: Julien Mairal (Department of Statistics, University of California, Berkeley) and Jean Ponce (Willow project-team, INRIA Paris-Rocquencourt).

Sparse coding, which is the decomposition of a vector using only a few basis elements, is widely used in machine learning and image processing. The basis set, also called dictionary, is learned to adapt to specific data. This approach has proven to be very effective in many image processing tasks. Traditionally, the dictionary is an unstructured "flat" set of atoms. In this work, we study structured dictionaries which are obtained from an epitome, or a set of epitomes. The epitome is itself a small image, and the atoms are all the patches of a chosen size inside this image. This considerably reduces the number of parameters to learn and provides sparse image decompositions with shift invariance properties. We propose a new formulation and an algorithm for learning the structured dictionaries associated with epitomes, and illustrate their use in image denoising tasks. This work has resulted in a CVPR’11 publication [11].

6.18. Dictionary Learning for Deblurring and Digital Zoom

**Participant:** Florent Couzinié-Devy.
The paper [23] proposes a novel approach to image deblurring and digital zooming using sparse local models of image appearance. These models, where small image patches are represented as linear combinations of a few elements drawn from some large set (dictionary) of candidates, have proven well adapted to several image restoration tasks. A key to their success has been to learn dictionaries adapted to the reconstruction of small image patches [35]. In contrast, recent works have proposed instead to learn dictionaries which are not only adapted to data reconstruction, but also tuned for a specific task [34]. We introduce here such an approach to deblurring and digital zoom, using pairs of blurry/sharp (or low-/high-resolution) images for training, as well as an effective stochastic gradient algorithm for solving the corresponding optimization task. Although this learning problem is not convex, once the dictionaries have been learned, the sharp/high-resolution image can be recovered via convex optimization at test time. Experiments with synthetic and real data demonstrate the effectiveness of the proposed approach, leading to state-of-the-art performance for non-blind image deblurring and digital zoom.

6.19. Robust linear least squares regression

Participant: Jean-Yves Audibert.
Collaboration with: Olivier Catoni (École Normale Supérieure, CNRS and INRIA Paris-Rocquencourt, Classic project-team)

In [26], we consider the problem of robustly predicting as well as the best linear combination of \(d\) given functions in least squares regression, and variants of this problem including constraints on the parameters of the linear combination. For the ridge estimator and the ordinary least squares estimator, and their variants, we provide new risk bounds of order \(d/n\) without logarithmic factor unlike some standard results, where \(n\) is the size of the training data. We also provide a new estimator with better deviations in presence of heavy-tailed noise. It is based on truncating differences of losses in a min-max framework and satisfies a \(d/n\) risk bound both in expectation and in deviations. The key common surprising factor of these results is the absence of exponential moment condition on the output distribution while achieving exponential deviations. All risk bounds are obtained through a PAC-Bayesian analysis on truncated differences of losses. Experimental results strongly back up our truncated min-max estimator. This work is to appear in the Annals of Statistics in 2012.

6.20. Semantic hierarchies for image annotation

Participant: Jean-Yves Audibert.
Collaboration with: Anne-Marie Tousch (École des Ponts and ONERA) and Stéphane Herbin (ONERA)

In the survey [6], we argue that using structured vocabularies is capital to the success of image annotation. We analyze literature on image annotation uses and user needs, and we stress the need for automatic annotation. We briefly expose the difficulties posed to machines for this task and how it relates to controlled vocabularies. We survey contributions in the field showing how structures are introduced. First we present studies that use unstructured vocabulary, focusing on those introducing links between categories or between features. Then we review work using structured vocabularies as an input and we analyze how the structure is exploited.

6.21. Deviations of Stochastic Bandit Regret

Participant: Jean-Yves Audibert.
Collaboration with: Antoine Salomon (École des Ponts)

The work [19] studies the deviations of the regret in a stochastic multi-armed bandit problem. When the total number of plays \(n\) is known beforehand by the agent, previous works exhibit a policy such that with probability at least \(1 - 1/n\), the regret of the policy is of order \(\log n\). They have also shown that such a property is not shared by the popular UCB1 policy. This work first answers an open question; it extends this negative result to any anytime policy. The second contribution of this paper is to design anytime robust policies for specific multi-armed bandit problems in which some restrictions are put on the set of possible distributions of the different arms.
6.22. Minimax Policies for Combinatorial Prediction Games

**Participant:** Jean-Yves Audibert.

Collaboration with: Sébastien Bubeck (Centre de Recerca Matematica of Barcelona) and Gabor Lugosi (ICREA and Pompeu Fabra University)

In [8], we address the online linear optimization problem when the actions of the forecaster are represented by binary vectors. Our goal is to understand the magnitude of the minimax regret for the worst possible set of actions. We study the problem under three different assumptions for the feedback: full information, and the partial information models of the so-called “semi-bandit”, and “bandit” problems. We consider both $L_\infty$, and $L_2$-type of restrictions for the losses assigned by the adversary. We formulate a general strategy using Bregman projections on top of a potential-based gradient descent, which generalizes the ones studied in numerous recent works. We provide simple proofs that recover most of the previous results. We propose new upper bounds for the semi-bandit game. Moreover we derive lower bounds for all three feedback assumptions. With the only exception of the bandit game, the upper and lower bounds are tight, up to a constant factor. Finally, we answer an open question raised by W. M. Koolen, M. K. Warmuth, and J. Kivinen by showing that the exponentially weighted average forecaster is suboptimal against $L_\infty$ adversaries.


**Participant:** Francis Bach.

Collaboration with: Eric Moulines (Telecom ParisTech)

In [10], we consider the minimization of a convex objective function defined on a Hilbert space, which is only available through unbiased estimates of its gradients. This problem includes standard machine learning algorithms such as kernel logistic regression and least-squares regression, and is commonly referred to as a stochastic approximation problem in the operations research community. We provide a non-asymptotic analysis of the convergence of two well-known algorithms, stochastic gradient descent (a.k.a. Robbins-Monro algorithm) as well as a simple modification where iterates are averaged (a.k.a. Polyak-Ruppert averaging). Our analysis suggests that a learning rate proportional to the inverse of the number of iterations, while leading to the optimal convergence rate in the strongly convex case, is not robust to the lack of strong convexity or the setting of the proportionality constant. This situation is remedied when using slower decays together with averaging, robustly leading to the optimal rate of convergence. We illustrate our theoretical results with simulations on synthetic and standard datasets.
5. New Results

5.1. Proliferation dynamics and its control

5.1.1. Cell division dynamics in structured cell populations

Participants: José Luis Avila Alonso [DISCO project-team, INRIA Saclay IdF], Annabelle Ballesta, Houda Benjelloun [INSA Rouen], Frédérique Billy, Frédéric Bonnans [Commands project-team, INRIA Saclay IdF], Catherine Bonnet [DISCO project-team, INRIA Saclay IdF], Jean Clairambault, Luna Dimitrio, Marie Doumic-Jaffreit, Xavier Dupuis [Commands project-team], Olivier Fercoq [MaxPlus project-team, INRIA Saclay IdF], Stéphane Gaubert [MaxPlus project-team, INRIA Saclay IdF], Germain Gillet [IBCP, Université Cl. Bernard Lyon 1], Philippe Gonzalo [IBCP, Université Cl. Bernard Lyon 1], Pierre Hirsch [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Thomas Lepoutre [now in DRACULA project-team, INRIA Rhône-Alpes, Lyon], Jonathan Lopez [IBCP, Université Cl. Bernard Lyon 1], Pierre Magal [University Bordeaux II], Anna Marciniak-Czochra [Institute of Applied Mathematics, Universität Heidelberg], Jean-Pierre Marie [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Faten Merhi [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Roberto Natalini [IAC-CNR, Università Sapienza, Rome], Silviu Niculescu [DISCO project-team, INRIA Saclay IdF], Hitay Özbay [Bilkent University, Ankara, Turkey], Benoît Perthame, Ruoping Tang [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Vitaly Volpert [CNRS Lyon, UMR5208, Camille Jordan Institute, Lyon], Jorge Zubelli [IMPA, Rio de Janeiro].

1. Transition kernels in a McKendrick model of the cell division cycle. A focus has been set on transitions between phases of the cell division cycle. The underlying biological question is: “Is desynchronisation between cells in proliferating cell populations a hallmark of cancer?” It has been considered by relating in a natural way transition kernels with the probability density functions of transition times in the cell population. It has been shown—which was expected, but never proved to our knowledge so far—that the more desynchronised cells are with respect to cell cycle phase transitions, the higher is the growth exponent of the cell population [48], otherwise said: desynchronised cell populations grow faster. This has been proven when transition kernels are time-independent, i.e., when no external control is exerted on transitions. The same question is currently experimentally investigated by our biologist partners in the European network ERASysBio+ C5Sys, coordinated by F. Lévi (Villejuif) and D. Rand (Warwick). Simulations using experimentally identified transition kernels in proliferating cell cultures controlled by theoretical time-dependent (circadian) control functions have verified the relevance of this mathematical result for theoretical cancer treatment optimisation (cf. infra “Periodic (circadian) control of cell proliferation in a theoretical model of the McKendrick type”).

2. Modelling haematopoiesis with applications to AML. The stability of a delay system based on a PDE model designed by M. Adimy and F. Crauste, structured by a discrete differentiation variable and multiple delays, with applications to Acute Myeloblastic Leukaemia (AML, clinical advisers: J.-P. Marie and P. Hirsch; technical adviser: RP Tang) is studied with possible therapeutic implications [36]. This model is currently experimentally investigated, with the aim to identify its parameters in leukaemic cells, in the DIGITEO project ALMA (cf. infra “DIGITEO and Cancérople IdF” in “Regional initiatives”), coordinated by C. Bonnet (DISCO team, INRIA Saclay IdF) and in the recently launched DIGITEO project ALMA2 (coordinated by J. Clairambault), that takes over the combined experimental-modelling activity in ALMA. Two INRIA postdocs, F. Merhi (in ALMA, 2010-2011) and A. Ballesta (in ALMA2, 2011-2013) have been devoted to this task. From a theoretical point of view, the Adimy-Crauste model has been modified so as a) to include quick
self-renewal of cells in each stage of maturation and b) to represent each phase of the proliferating compartment (i.e., $G_1$, $S$, $G_2$ and $M$) separately. For the time being, only the $M$ phase is supposed to have a fixed time duration as it is generally admitted that the short time (typically half an hour if the total proliferating phase duration is normalised to 24 hours) necessary to perform mitosis is hardly submitted to any variation.

In a complementary manner, a new model for cell differentiation was introduced and analysed in [17], in collaboration with A. Marciniak and J.P. Zubelli. It assumed that differentiation of progenitor cells is a continuous process. From the mathematical point of view, it is based on partial differential equations of transport type. Specifically, it consists of a structured population equation with a nonlinear feedback loop. This models the signaling process due to cytokines, which regulate the differentiation and proliferation process. We compared the continuous model to its discrete counterpart, a multicompartmental model of a discrete collection of cell subpopulations recently proposed by Marciniak-Czochra et al. [Stem Cells Dev., 18 (2009), pp. 377–386] to investigate the dynamics of the hematopoietic system. We obtained uniform bounds for the solutions, characterized steady state solutions, and analyzed their linearized stability. We showed how persistence or extinction might occur according to values of parameters that characterize the stem cells’ self-renewal. We also performed numerical simulations and discuss the qualitative behavior of the continuous model vis-à-vis the discrete one.

3. **Hybrid models**

Systems combining PDEs and discrete representations in hybrid models, with applications to cancer growth and therapy, in particular for AML, are the object of study of the ANR program *Bimod*, coordinated by V. Volpert (Lyon), associating CNRS (V. Volpert, Lyon), Bordeaux II University (P. Magal) and the Bang project-team.

4. **Molecular model of apoptosis.**

With G. Gillet (prof. at IBCP/Lyon), we have designed a mathematical ODE model for the mitochondrial pathway of apoptosis, focused on the early phase of apoptosis (before the cytochrome C release). We have validated it with experimental data carried out in G. Gillet’s lab and applied it to propose new therapeutic strategies against cancer. This work has led to a nearly submitted article [47].

5. **Molecular model of the activity of the p53 protein.** Following her first year of PhD in Rome with R. Natalini, working on cytoplasmic transport along microtubules presented in [38], L. Dimitrio has begun her third PhD year, going on studying at INRIA nucleocytoplasmic transport with applications to p53 activity. Her PhD thesis work is supervised in co-tutela between Sapienza University in Rome (R. Natalini) and INRIA (J. Clairambault). The protein p53 plays a capital part as “guardian of the genome”, arresting the cell cycle and launching cell apoptosis or DNA repair in case of DNA damage. Results expected from this newly developed theme will provide a rational link between molecular pharmacokinetics-pharmacodynamics (cf. infra) of anticancer drugs and modelling of the cell division cycle in proliferating cell populations. L. Dimitrio has presented her ongoing work in different meetings in France and in Italy, and a paper in preparation will be submitted in 2012.

5.1.2. **Physiological and pharmacological control of cell proliferation**

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

1. **Periodic (circadian) control of cell proliferation in a theoretical model of the McKendrick type.** The impact of a periodic control exerted on cell cycle phase transitions has continued to be studied [16] with the collaboration of S. Gaubert (MaxPlus INRIA project-team, Saclay IdF) and T. Lepoutre (Dracula INRIA project-team, Lyon) and is currently investigated experimentally in the new C5Sys European network (cf. supra “Transition kernels in a McKendrick model of the cell division cycle”
and “). Thanks to the work of Frédérique Billy (Postdoc in Bang) and Olivier Fercoq (PhD student in MaxPlus), together with permanent members of Bang, Dracula and MaxPlus teams, it has led to three publications \[37\], \[39\], \[48\].

2. **Intracellular pharmacokinetic-pharmacodynamic (PK-PD) models for anticancer drugs.** This theme is actively worked on in collaboration, mainly with the teams of F. Lévi and J.-P. Marie (cf. supra “Transition kernels in a McKendrick model of the cell division cycle” and “Modelling haematopoiesis with applications to AML”). After a PK-PD model for 5-FU with folinic acid \[86\], it has led for the anticancer drug Irinotecan, the main object of A. Ballesta’s PhD thesis \[1\], to an article published in PLoS Computational Biology \[8\], reporting a combined modelling and experimental approach to the effects of a combination of mathematical modelling and experimentation in cell cultures, and to another one \[7\], focusing on drug delivery optimisation.

3. **Whole body physiologically based model of anticancer drug pharmacokinetics.** This theme has also been studied in A. Ballesta’s PhD thesis. The use of identification, in genetically different laboratory mouse strains, of parameters characterising an ODE model of the action of Irinotecan (cf. supra “Intracellular pharmacokinetic-pharmacodynamic (PK-PD) models for anticancer drugs”) in cell cultures, transposed at the whole-body level, has been designed as a proof of concept for individual adaptation of drug delivery in the context of (future) personalised medicine, a perspective sketched in \[15\] and among other collaborative contexts linking mathematics and medicine in \[14\], \[3\].

5.1.3. **Optimisation of cancer chemotherapy**

**Participants:** Annabelle Ballesta, Frédérique Billy, Frédéric Bonnans [Commands project-team], Jean Clairambault, Sandrine Duloung [INSERM Villejuif (U 776)], Xavier Dupuis [Commands project-team], Olivier Fercoq [MaxPlus project-team], Stéphane Gaubert [MaxPlus project-team], Thomas Lepoutre [Dracula project-team], Alexander Lorz, Francis Lévi [INSERM U 776, Villejuif], Michael Hochberg [ISEM, CNRS, Montpellier], Benoît Perthame.

Optimising cancer chemotherapy, especially chronotherapy, is the final aim of the activities mentioned above. This has been lately discussed in \[16\] and also in works involving the C5Sys network \[37\], \[48\], and in the more general review \[39\]. Until now had been taken into account as constraints in optimisation strategies only the unwanted toxic side effects of anticancer drugs on healthy cells. More recently, another issue of anticancer treatment has been considered, namely the different mechanisms of resistance to drugs in cancer cells. This has led to include the effect of ABC transporters (active efflux pumps, as is the P-glycoprotein) in the intracellular PK-PD models mentioned above \[86\], in A. Ballesta’s PhD joint work with F. Lévi’s team \[1\], \[8\], \[7\], and to a perspective paper \[15\].

In project is also the use of methods of optimal control developed by the Commands project-team (F. Bonnans, X. Dupuis) to optimise therapies in the treatment of Acute Myeloblastic Leukaemia (AML, cf. supra “Modelling haematopoiesis with applications to AML”).

Another way to represent and overcome drug resistance in cancer from a cell Darwinian point of view using concepts of adaptive dynamics in proliferating cell populations is also currently being investigated along the line of other recent works \[28\] and is currently developed within the multidisciplinary GDR DarEvCan coordinated by M. Hochberg, Montpellier (cf. infra “GDR DarEvCan” in “National initiatives”) and in a proposed ANR project also coordinated by M. Hochberg.

An open question that should have therapeutic implications consists of the interrogation: Is the emergence of drug resistance in cell populations a genetic (resulting from mutations at mitosis) or an epigenetic phenomenon (resulting from amplification of physiological mechanisms, such as ABC transport, which has nothing to do with genetic mutations)? And is it a reversible or irreversible phenomenon? These questions will be studied both theoretically and experimentally within the DarEvCan consortium and could result in new developments in the so-called Darwinian medicine.
5.1.4. Protein polymerisation and application to amyloid diseases (ANR grant TOPPAZ)

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

With H. Rezaei, a new and very complete PDE model for protein polymerisation has been designed. Following F. Charles’ work, A. Ballesta has applied this model to Huntington’s disease (PolyQ expansion) and compared it with its ODE counterpart, leading to a better understanding of the leading mechanisms responsible for PolyQ fibrillation. This part is nearly submitted.

The eigenvalue problem playing a major role in the representation of Prion proliferation dynamics and, in a more general way, of many fragmentation-coalescence phenomena, the article [11] investigated the dependency of the principal eigenvector and eigenvalue upon its parameters. We exhibited possible non-monotonic dependency on the parameters, conversely to what would have been conjectured on the basis of some simple cases.

5.1.5. Inverse problem in growth-fragmentation equations

**Participants:** Marie Doumic-Jauffret, Marc Hoffmann [ENSAE], Patricia Reynaud [CNRS, Nice Univ.], Vincent Rivoirard [Paris IX Univ.], Léon Matar Tine [SIMPAF project-team, INRIA Lille Nord-Europe].

In collaboration with statisticians (M. Hoffman, Professor at Université de Marne-la-Vallée, V. Rivoirard, MC at Université d’Orsay, and P. Reynaud, CR CNRS at Université de Nice), in [18] we have explored a statistical viewpoint on the cell division problem. In contrast to a deterministic inverse problem approach, we take the perspective of statistical inference. By estimating statistically each term of the eigenvalue problem and by suitably inverting a certain linear operator, we are able to construct an estimator of the division rate that achieves the same optimal error bound as in related deterministic inverse problems. Our procedure relies on kernel methods with automatic bandwidth selection. It is inspired by model selection and recent results of Goldenschluger and Lepski. This work is accepted in SIAM J. Num. Anal..

With L. Matar Tine, in [53] we have generalized the inverse techniques proposed in [88], [67] and [66], in order to adapt them to general fragmentation kernels and growth speeds. The potential applications of this problem are numerous, ranging from polymerisation processes to the cell division cycle. This work is submitted.

5.2. Tissue growth, regeneration and cell movements

5.2.1. Chemotaxis, self-organisation of cell communities

**Participants:** Nikolaos Bournaveas [Univ. Edinburgh], Axel Buguin [UPMC, Institut Curie], Vincent Calvez [ENS Lyon], François James [univ. Orléans], Alexander Lorz, Grégoire Nadin [UPMC], Benoît Perthame, Jonathan Saragosti [Institut Curie], Pascal Silberzan [Institut Curie], Min Tang [SJTU], Nicolas Vauchelet.

We have continued our analysis and simulation of models for large bacterial communities and more generally cells self-organisation as initiated several years ago [12], [31]. This is a rich domain because on the one hand several Partial Differential Equations arise, parabolic models, kinetic equations, hyperbolic systems and on the other hand complex patterns occur that are a sign of the complex underlying dynamics.

In their article [65], Y. Dolak and C. Schmeiser have proposed a mathematical model describing the individual behavior of bacteria responding to a chemical substance. Numerical simulations of this kinetic model have been investigated in [93]. In a macroscopic level of description, we perform a hydrodynamical limit which leads to an aggregation model, for which regular solutions blows up in final time [82]. The rigorous study of the behavior of the aggregates relies on a careful analysis of some non-linear scalar conservation laws in the framework of measures [24] and has been investigated in [55].
With the group of P. Silberzan in Curie Institute, in 2010 we have given an explanation of traveling bands first observed by Adler in the 80’s for *E. coli* in microchannels. In a continuation of this work, based on analysis of individual trajectories, we have shown directional persistence which improves the efficiency of collective migration. Kinetic models with tumbling kernels that keep memory of the incoming velocity are able to reproduce accurately the wave parameters [35].

Computing effectively traveling wave can be difficult in unstable cases. An algorithm is proposed in [29] for the NonLocalFisher equation which catches the traveling wave and not the more stable pulsating wave.

A. Lorz has studied a system consisting of the elliptic-parabolic Keller–Segel equations coupled to Stokes equations by transport and gravitational forcing. We show global-in-time existence of solutions for small initial mass in 2D. In 3D we establish global existence assuming that the initial $L^{3/2}$-norm is small. Moreover, we give numerical evidence that for this extension of the Keller–Segel system in 2D, solutions exist with mass above $8\pi$, which is the critical mass for the system without fluid. The model is written as

$$
\begin{align*}
    u \cdot \nabla c &= \Delta c + n - a_1 c, \\
    n_t + u \cdot \nabla n &= \Delta n - \nabla \cdot (\chi n \nabla c), \\
    a_2 u_t + \nabla P - \eta \Delta u + n \nabla \phi &= 0, \\
    \nabla \cdot u &= 0.
\end{align*}
$$

(2)

Here $c$ denotes the concentration of a chemical, $n$ a cell density and $u$ a fluid velocity field described by Stokes equations. The fluid couples to $n$ and $c$ through transport and gravitational forcing modelled by $\nabla \phi$. The pressure $P$ can be seen as the Lagrange multiplier enforcing the incompressibility constraint. The chemical $c$ diffuses, it is produced by the cells and it degrades. The cell density diffuses and it moves in the direction

Figure 1. Dynamics of the swarmers and supporters $n + f$ concentration (left) and of the surfactin $S$ (right) solving system
of the chemical gradient. The constant $a_1 \geq 0$ measures self-degradation of the chemical and the constants $a_2 \geq 0$, $\eta > 0$ determine the evolution undergone by $u$.

5.2.2. Single cell-based models of tumour growth, tissue regeneration, embryonic development

Participants: Annabelle Ballesta, Gregory Batt [CONTRAINTES project-team], François Bertaux, Chadha Chettaoui, Ibrahim Cheddadi, Dirk Drasdo, Adrian Friebel, Rolf Gebhardt [Univ. of Leipzig, Germany], Adriano Henney [Director Virtual Liver Network and VLN consortium], Jan G. Hengstler [Leibniz Research Center, Dortmund, Germany and CANCERSYS consortium], Stefan Höhme, Elmar Heinzle [University of Saarbrücken and NOTOX consortium], Isabelle Hue [INRA], Nick Jagiella, Ursula Klingmüller [German Cancer Center, Heidelberg and LungSys Consortium], Axel Krinner, Emanuele Leoncini, Johannes Neitsch, Benoît Perthame, Ignacio Ramis-Conde, Luc Soler [ IRCAD, Coordinator EU-project PASSPORT and PASSPORT consortium], Irène Vignon-Clémentel [REO project-team], Juhui Wang [INRA], William Weens.

Structure formation in tissues as well as malfunctions on the multi-cellular level are inherently of multi-scale nature. Modifications on the molecular level by intrinsic or extrinsic factors affect the architecture and function on the multi-cellular tissue level. Much of the current research so far focuses on the analysis of intracellular pathways, genetic and metabolic regulation on the intracellular scale and on continuum equations for local densities of cells to capture multi-cellular objects on large spatial scales but only recently have increasing efforts been made at the interface between these two: individual cell based models (IBMs) which permit to include the molecular information on the one hand and to extrapolate to the multi-cellular tissue level on the other hand and hybrid models that combine continuum with individual-based models for different components.

In order to fill the existing gap we have studied intracellular regulation networks \cite{87, 72}, multi-scale IBMs where intracellular regulation and differentiation was explicitly represented within each individual cell \cite{90, 85, 91}, lattice-free IBMs \cite{69} and continuum models that can capture their large scale behaviour \cite{63}, and cellular automaton (CA) models where each lattice site can be occupied either by at most one cell \cite{61} or by many cells \cite{89, 71} and their corresponding continuum equation \cite{68}. Moreover, for a simple, but for rigorous coarse graining not accessible, growth situation we were able to obtain quantitatively matching results with continuum and individual-cell-based models without any fit parameter.
Besides the methodical aspects we focus on a number of applications:

- Unstructured cell populations growing in a monolayer with free border [69], [74], or constraint by the presence of a granular or cellular embedding medium [19].

- Multicellular spheroids in liquid suspension [69], [70], and embedding granular or cellular matter [19]. For non-small-lung-cancer cell lines growing as multi-cellular spheroids, we could, starting with a complete parameterisation of the model by labelling experiments, simultaneously explain the proliferation, apoptosis, extracellular matrix and growth pattern of multi-cellular spheroids under different nutrient conditions within one consistent mathematical model.

- Vascular tumour growth.

- Regulatory and evolutionary aspects in tumour growth [81], [78], recently with resolution of intracellular signal transduction pathway variants found in normal vs. malignant cells [33].

- Cell differentiation and lineage commitment of mesenchymal stem cells [85], [73]. In our earlier work we have established a model of cell aging for in-vitro cultured stem cell populations. Stem cell concepts developed earlier [85], [73] have been extended to include cell aging [83]. By this extension it is possible to explain the clonal heterogeneity that was not captured by the previous model. The cell age was coupled with the generation number. It is published in ref. [84] and [56].

- Complex tissue architectures in regenerative tissues, particularly in the liver.

Examples are:

- Regeneration of liver lobules after toxic damage [80], [79], [76], [40], [41], [42] within the German BMBF-funded network “Systems Biology of the Hepatocyte”). As extension of this project we linked the regeneration of liver architecture after toxic damage to a model of ammonia detoxification by the individual hepatocyte and the liver as a proof of concept to study the link between architecture and function. The comparison of experimental findings by our collaboration partners with our model results suggests that the detoxification during regeneration after drug-induced damage is mainly determined by the total population size of healthy hepatocytes (Fig. 2). Adjustments of enzymatic activities of the individual hepatocytes seem to have only minor effects.

- Liver regeneration after partial hepatectomy ([44]). Based on the work on regeneration of a liver after toxic damage where we focused on a single liver lobule, we within the EU project
CANCERSYS set up a model on liver regeneration after partial hepatectomy enabling us to model up to the whole liver lobe scale of a mouse, 4 cells thick. This models permits to bridge the gap between the single-cell-model scale and the whole-liver organ scale. Calibrating this model with mouse data we were able to predict the proliferation pattern in pig as a proof of principle that modelling can be used to bridge the gap between different animals. Experiments performed so far confirm the prediction. This is a fundamental issue as it is a longstanding unsolved question in how far experiments in animal models can be used to predict therapeutic responses in the Human. We also expanded the software towards whole liver lobe bright field image analysis. Being able to use a mathematical model calibrated with data from a model animal, for example, mouse to predict tissue organisation processes in another animal, for example, human opens new possibilities to assess drug toxicity [43].

- The fundamental objective is to move towards modelling from the molecular up to the whole organ scale. A conceptual framework for this was presented in ref. [23].
- Cancer development in an environment of granular particles and cells. For the first part we studied how an embedding medium such as granular particles or cells modify the spatial and temporal growth pattern of expanding cell populations [19]. The model could explain the growth kinetics found by Helmlinger et. al. (1997) for growing tumor spheroids embedded in agarose gel. If the friction between the embedding objects and the environment is larger than the friction between the growing clone and its environment we found a fingering instability reminiscent of a Saffman-Taylor instability observed in a Hele-Shaw cell. We systematically studied which model parameters promote the instability. The motivation for this project was to analyse in how far invading tumour fronts can be explained by physical mechanisms alone as in some tumour phenotypes invasive fronts can be observed.

In order to extend this project towards carcinoma development mechanisms of liver cancer development in mice were studied [42]. Within systematic sensitivity analyses we were able to identify parameters explaining the experimentally found tumor phenotypes (Fig. 3). The critical parameters were stiffness of sinusoids, the micro-blood vessels within the liver lobules, tumor cell - sinusoidal adhesion, tumor cell polarity, and the ability of tumour cells to digest neighbouring vessels. As part of the project, in close collaboration with experimental partners, critical differences between liver in normal and transgenic mice have been studied by image analysis [62].

- Synthetic biology. By multi-scale simulations including intracellular pathways in our single-cell-based simulation framework we were able to mimic conditions under which tissue homeostasis and tissue location could be achieved in monolayer culture.

The applications are guided by quantitative comparisons to experimental data either from published knowledge or - in most cases - generated by experimental partners. One main focus is on the understanding of mechanisms that control the growth dynamics and growth phenotypes of multi-cellular systems and use these later to predict and optimise therapy or biotechnological growth processes.

The adjustment of the models developed to applications requires data analysis both, of molecular data such as gene expression profiles and of image data such as spatial-temporal growth pattern. For this purpose we recently considered the geometric and topological measures to quantify tumour shapes [92], and developed an image processing chain to quantitatively analyse liver regeneration processes in liver lobules [79], [76], [44], [40], [41]. As a further step we published executables and descriptions of important elements of our code to spread our model as it turns out that agent-based cell modelling enjoys increasing interest in different communities (engineering, mathematical biology, systems biology, physics) [81]. Current directions moreover include a stronger focus on models of in-vivo systems (within the German medical systems biology consortium “LungSys” (lung cancer treatment); and within the EU-network “CancerSys” (cancerogenesis in liver)). Within LungSys we recently developed a realistic 2D and 3D spatial temporal model of blood flow in xenografts to compare to DCE MRI images visualising the tumour perfusion. Modelling cancer development requires to take into account invasion, mutations and angiogenesis, three hallmarks of cancer and of linking the molecular to the multicellular scale [71]. Moreover, we extend the topic of liver regeneration to regeneration
after partial hepatectomy (within the EU-project “Passport”), and extend our modelling activities to understand early embryonic development (Trophoblast development, collaboration with INRA).

Almost each of our projects is in close collaboration with experimental partners within grant projects performing experiments to permit parameterisation and validation of our models.

5.3. Modeling in computational neurosciences

Participants: Maria Caceres [Univ. Granada], Jose Carrillo [ICREA Barcelona], Benoît Perthame, Jonathan Touboul.

Networks of interacting neurons can be well described by nonlinear PDEs like the Noisy Integrate and Fire model. These are Fokker-Planck-Kolmogorov equations on the probability density of neurons, the main parameters in the model being the connectivity of the network and the noise. In [10], we analyse several aspects of the NNLIF model: the number of steady states, a priori estimates, blow-up issues and convergence toward equilibrium in the linear case. In particular, for excitatory networks, blow-up always occurs for initial data concentrated close to the firing potential. These results show how critical is the balance between noise and excitatory/inhibitory interactions to the connectivity parameter.

At a larger scale, neurons form large-scale spatially extended populations receiving similar input and interconnected in a specific way. Each neuron receives noisy inputs, and as such their membrane potential is adequately described as the solution of stochastic network equations. In [58] we study the asymptotic regimes of such spatially extended networks with delays and obtain a complex mean-field equation the dynamics of which is analyze in [57]. We observe that noise induces transitions from stationary spatially homogeneous solutions to oscillatory solutions, and the transition is characterized by chaotic Turing patterns of activity.

5.4. Free surface geophysical flows

Participants: Emmanuel Audusse [LAGA - Université Paris 13, Institut Galilée], Sakina Ayata, Anne-Céline Boulanger, Marie-Odile Bristeau, Benoît Perthame, Jacques Sainte-Marie [CETMEF and MACS project-team].

We are involved in research concerning the numerical simulation of free surface geophysical flows such as rivers, lakes, coastal areas and also overland flows. Many applications related to environmental problems are concerned: floodings, dam breaks, swell, transport and diffusion of pollutants, water quality, upwellings, sustainability of aquatic ecosystems, ...

The basic model for these problems is the 3D free surface Navier-Stokes system leading to a 3D solver [64] with a moving mesh. However for efficiency reasons, vertically averaged models such as the Saint-Venant system [75] are often used.

The Saint-Venant equations are deduced of the Navier-Stokes system with two main assumptions:

- the pressure is hydrostatic,
- the horizontal velocity is represented by its average.

We have developed extensions of the Saint-Venant system where the basic Saint-Venant solver [60] is still used and, in that way, the robustness, the efficiency and the easiness to treat the free surface are preserved while the domain of validity is larger.

In these extensions, we relax the two above assumptions. Actually, we have derived a non-hydrostatic shallow water model and a multilayer Saint-Venant system.

We have coupled the hydrodynamics of free surface flows with other phenomena such as biology (phytoplankton culture) or erosion.
Figure 4. Map of Japan with the seism epicentre and the DART buoys 21418 and 21413.

Figure 5. Free surface elevation of the sea, comparison between the recorded data by the buoys 21418 and 21413 and the simulation obtained with our 3d Navier-Stokes code.
5.4.1. Hydrodynamics and biology coupling

Cultivating oleaginous microalgae in specific culturing devices is seen as a potential source of biofuel for the future. The complexity of this process coupling non-linear biological activity to hydrodynamics makes the optimization problem very delicate. The large amount of parameters to be taken into account paves the way for a useful mathematical modeling. Due to the high heterogeneity of raceways along the depth dimension regarding temperature, light intensity or nutrients availability, we adopt a multilayer approach for hydrodynamics and biology. For hydrodynamics, we use a multilayer Saint-Venant model that allows mass exchanges, forced by a simplified representation of the paddlewheel. Then, starting from an improved Droop model that includes light effect on algae growth, we derive a similar multilayer system for the biological part. A kinetic interpretation of the whole system results in an efficient numerical scheme. We show through numerical simulations in two dimensions that our approach is capable of discriminating between situations of moving water or calm pond and show the influence of light intensity. Moreover, in this paper [49] we exhibit that a posteriori treatment of our velocity fields can provide Lagrangian trajectories which are of great interest to assess the actual light pattern perceived by the algal cells and therefore understand its impact on the cell factory.

![Figure 6. Trajectories of three particles during the simulations. In every figure, the large curve represents the water height at the middle of the pool. The other plot is the height of a given particle through time. The algae undergo sudden changes of depth every time it meets the wheel.](image)

5.4.2. Analytical solutions for the free surface hydrostatic Euler equations

In this paper [50] we propose a large set of analytical solutions for the hydrostatic incompressible Euler system in 2d and 3d. These solutions mainly concern free surface flows but flows with partially free surface or in a deformable pipe are also considered. These analytical solutions that can admit entropic shocks can be especially useful for the validation of numerical schemes.

5.4.3. Phytoplankton growth in marine ecosystem

Four different phytoplankton growth models have been implemented. The simplest model assumes constant chlorophyll/carbon and carbon/nitrogen ratios. The more complex ones take into account photoadaptation through a variable chlorophyll/carbon ratio and they also assume a variable carbon/nitrogen cellular quota.
(non-redfieldian stochiometry). These models have been coupled to a 1D ecosystem model at BATS, a station located in an oligotrophic area of the North-Western Atlantic Ocean. The different models have been calibrated from in situ data recorded at BATS using a micro-genetic algorithm to optimize the parameter values. The models with optimized parameters were then compared with each others. The results highlighted the necessity to take into account photoadaptation and variable cellular quotas to simulate the seasonal dynamics of chlorophyll and primary production in oligotrophic areas. They also demonstrated that the chlorophyll did not have to be represented by a prognostic variable and could be represented by a diagnostic variable instead, see [46].

5.4.4. Erosion processes: modelling and simulation

We are interested in the modelling of sediment transport phenomena. We mostly focus on bedload transport and we do not consider suspension sediment processes. We first propose a coupled numerical scheme for the classical Saint-Venant – Exner model. It is based on a relaxation approach and it works with all sediment flux function. We exhibit that this coupled approach is more stable than the splitting approach that is mostly used in industrial softwares. Then we derive an original three layers model in order to overcome the difficulties that are encountered when using the classical Exner approach and we write a related relaxation model, see [45].
6. New Results

6.1. New methods for data assimilation

Since the beginning, CLIME has been focused on new techniques for data assimilation. Last year’s focus of the methodological development was on the use of non-Gaussian approaches for inverse modeling, and the construction of a multiscale data assimilation methodology. Several methodological papers have now been published on these topics. This year, the applications of these methodologies are put forward in the inverse modeling section, although new theoretical developments have been added to these approaches. In addition, new topics have been addressed, such as the ensemble Kalman filter with a theory that puts the EnKF on safer grounds, the use of 4D-Var for the estimation of parameters in dispersion models, and the real-time data assimilation at urban scale.

6.1.1. Ensemble Kalman filtering without the intrinsic need for inflation

Participant: Marc Bocquet.

The main intrinsic source of error in the ensemble Kalman filter (EnKF) is sampling error. External sources of error, such as model error or deviations from Gaussianity, depend on the dynamical properties of the model. Sampling errors can lead to instability of the filter, which, as a consequence, often requires inflation and localization. The goal of this study is to derive an ensemble Kalman filter, which is less sensitive to sampling errors. A prior probability density function conditional on the forecast ensemble is derived using Bayesian principles. Even though this prior is built upon the assumption that the ensemble is Gaussian-distributed, it is different from the Gaussian probability density function defined by the empirical mean and the empirical error covariance matrix of the ensemble, which is implicitly used in traditional EnKFs. This new prior generates a new class of ensemble Kalman filters, called finite-size ensemble Kalman filter (EnKF-N). One deterministic variant, the finite-size ensemble transform Kalman filter (ETKF-N), is derived. It is tested on the Lorenz ’63 and Lorenz ’95 models. In this context, ETKF-N is shown to be stable without inflation for ensemble size greater than the model unstable subspace dimension, at the same numerical cost as the ensemble transform Kalman filter (ETKF). One variant of ETKF-N seems to systematically outperform the ETKF with optimally tuned inflation. However, it is shown that ETKF-N does not account for all sampling errors and necessitates localization like any EnKF, whenever the ensemble size is too small. In order to explore the need for inflation in this small ensemble size regime, a local version of the new class of filters is defined (LETKF-N) and tested on the Lorenz ’95 toy model. Whatever the size of the ensemble, the filter is stable. Its performance without inflation is slightly inferior to that of LETKF with optimally tuned inflation for small interval between updates, and superior to LETKF with optimally tuned inflation for large time interval between updates.

6.1.2. Parameter field estimation for atmospheric dispersion: Application to the Chernobyl accident using 4D-Var

Participant: Marc Bocquet.

Atmospheric chemistry and air quality numerical models are driven by uncertain forcing fields: emissions, boundary conditions, wind fields, vertical turbulent diffusivity, kinetic chemical rates, etc. Data assimilation can help to assess these parameters or fields of parameter. Because those parameters are often much more uncertain than the fields diagnosed in meteorology and oceanography, data assimilation is much more an inverse modeling challenge in this context. In this study, we experiment with these ideas by revisiting the Chernobyl accident dispersion event over Europe. We develop a fast four-dimensional variational scheme (4D-Var), which seems appropriate for the retrieval of large parameter fields from large observations sets, and for the retrieval of parameters that are non-linearly related to concentrations. The 4D-Var, and especially an approximate adjoint of the transport model, is tested and validated using several advection schemes that are
influential on the forward simulation as well as on the data assimilation results. Firstly, the inverse modeling system is applied to the assessment of the dry deposition parameters and of the wet deposition parameters. It is then applied to the retrieval of the emission field alone, to the joint optimization of removal process parameters and source parameters, and to the optimization of larger parameter fields, such as horizontal and vertical diffusivities, or dry deposition velocity field. The physical parameters used so far in the literature for the Chernobyl dispersion simulation are partly supported by the study. The crucial question of deciding whether such an inversion is merely a tuning of parameters, or a retrieval of physically meaningful quantities is discussed. Even though inversion of parameter fields may fail to determine physical values for the parameters, it achieves statistical adaptation that partially corrects for model errors, and, using the inverted parameter fields, leads to considerable improvement in the simulation scores.

![Figure 2. Validation of the parameter estimation approach using 4D-Var for dry deposition and wet scavenging parameters in the context of the Chernobyl dispersion event. The 4D-Var optimization and the brute force screening value approach lead to the same result.](image)

### 6.1.3. Real-time data assimilation

**Participants:** Vivien Mallet, Anne Tilloy, Fabien Brocheton [Numtech], David Poulet [Numtech], Cécile Honoré [Airparif], Édouard Debry [INERIS].

Based on Verdandi, Polyphemus and the “Urban Air Quality Analysis” software, real-time data assimilation was carried out at urban scale. The Best Linear Unbiased Estimator (BLUE) was computed for every hourly concentration map that the ADMS model computed. A posteriori tests were conducted over Clermont-Ferrand and Paris. We addressed the key issue of the covariance of the state error. The form of the error covariance between two points was determined based on the road network, considering the distance between points along the road and the distance of each point to the road. A few parameters (primarily two decorrelation lengths) were determined thanks to cross validation with several months of simulations and observations. The results showed strong improvements even at locations where no data was assimilated. The assimilation was carried out in the prototype “Votre Air” ([http://votreair.airparif.fr/](http://votreair.airparif.fr/)) for real-time air quality estimation over a part of Paris, which originally run in the context of the Futur-en-Seine festival and now runs operationally since June 2011.
At larger scale, the data assimilation library Verdandi was used to apply data assimilation (optimal interpolation) with the air quality model Chimere. This preliminary work will help INERIS to apply optimal interpolation in the operational platform Prev’air.

6.2. Inverse modeling

Many of this year’s studies have focused on inverse modeling, including the reconstruction of the Fukushima radionuclide source term. All were targeted to a particular application. However most of them include new methodological developments.

6.2.1. Estimation of errors in the inverse modeling of accidental release of atmospheric pollutant: Application to the reconstruction of the Fukushima Daiichi source term

Participants: Victor Winiarek, Marc Bocquet, Olivier Saunier [IRSN], Anne Mathieu [IRSN].

The aim of this research activity is the implementation of data assimilation methods, particularly inverse modeling methods, in the context of an accidental radiological release from a nuclear power plant and their application in the specific case of the Fukushima Daiichi accident. The particular methodological focus is the a posteriori estimation of the prior errors statistics. In the case of the Fukushima Daiichi accident, the number of available observations is small compared to the number of source parameters to retrieve and the reconstructed source is highly sensitive to the prior errors. That is the why they need to be well established and justified.

In this aim, three methods have been proposed: one method relies on an L-curve estimation technique, another one on the Desroziers’ iterative scheme and the last method, assumed to be the most robust, relies on the maximum likelihood principle (generalized to a non-Gaussian context).

These three methods have been applied to the reconstruction of cesium-137 and iodine-131 source terms from the Fukushima Daiichi accident. Because of the poor observability of the Fukushima Daiichi emissions, these methods provide lower-bounds for cesium-137 and iodine-131 reconstructed activities. Nevertheless, with the new method based on semi-Gaussian statistics for the background errors, these lower-bound estimates, $1.2 - 4.0 \times 10^{16}$ Bq for cesium-137, with an estimated standard deviation range of $15 - 20\%$, and $1.9 - 3.8 \times 10^{17}$ Bq for iodine-131, with an estimated standard deviation range of $5 - 10\%$, are of the same order of magnitude as those provided by the Japanese Nuclear and Industrial Safety Agency, and about 5 to 10 times less than the Chernobyl atmospheric releases.

6.2.2. Optimal Representation of Source-Sink Fluxes for Mesoscale Carbon Dioxide Inversion

Participants: Lin Wu, Marc Bocquet, Frédéric Chevallier [LSCE, CEA], Thomas Lauvaux [Department of Meteorology, Pennsylvania State University, USA], Peter Rayner [School of Earth Sciences, University of Melbourne, Australia], Ken Davis [Department of Meteorology, Pennsylvania State University, USA].

This study is an application of our previous theoretical developments on a consistent Bayesian multiscale formalism to optimally design control space (in which control variables are to be estimated). We construct the optimal adaptive representations of the surface fluxes for mesoscale carbon dioxide inversions. Such representations are taken from a large dictionary of adaptive multiscale grids. These optimal representations are obtained by maximizing the number of Degrees of Freedom for the Signal (DFS) that measures the information gain from observations to resolve the unknown fluxes. Consequently information from observations can be better propagated within the domain through these optimal representations.

The optimal representations are constructed using synthetic continuous hourly carbon dioxide concentration data in the context of the Ring 2 experiment in support of the North American Carbon Program Mid Continent Intensive (MCI). Compared with the regular grid at finest scale, optimal representations can have similar inversion performance with far fewer grid cells. For the Ring 2 network of eight towers, in most cases, the DFS value is relatively small compared to the number of observations $d$ (DFS/$d < 20\%$). In this multiscale setting, scale-dependent errors due to the aggregation of flux variables are identified and explicitly formulated for more reliable inversions. It is recommended that this aggregation error should be taken into account, especially when the correlations in the errors of a priori fluxes are physically unrealistic. The optimal multiscale grids allow to adaptively mitigate the aggregation error.
6.3. Monitoring network design

In this section, we report studies that are related to the evaluation of monitoring network and to new monitoring strategies. As opposed to last year’s report, they may not be strictly addressing optimal network design.

6.3.1. Eyjafjallajökull ash concentrations derived from both Lidar and modeling

**Participants:** Patrick Chazette [LSCE], Marc Bocquet, Philippe Royer [LSCE], Victor Winiarek, Jean-Chistophe Raut [ATMOS], Philippe Labazuy [OPG/LMV], Mathieu Gouhier [OPG/LMV], Mélody Lardier [LEOSPHERE], Jean-Pierre Cariou [LEOSPHERE].

Following the eruption of the Icelandic volcano Eyjafjallajökull on April 14 2010, ground-based N2-Raman lidar (GBL) measurements were used to trace the temporal evolution of the ash plume from April 16 to April 20 2010, above the southwestern suburb of Paris. The nighttime overpass of the Cloud-Aerosol LIdar with Orthogonal Polarization onboard Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation satellite (CALIPSO/CALIOP), on April 17 2010, was an opportunity to complement GBL observations. The plume shape retrieved from GBL has been used to assess the size range of the particles size. The lidar-derived aerosol mass concentrations (PM) have been compared with model-derived PM concentrations held in the Eulerian model Polair3D transport model, driven by a source term inferred from the SEVIRI sensor on board the satellite Meteosat. The consistency between model and ground-based wind lidar and CALIOP observations has been checked. The spatial and temporal structures of the ash plume as estimated by each instrument and by the Polair3D simulations are in good agreement.

6.3.2. Potential of the International Monitoring System radionuclide network for inverse modeling

**Participants:** Mohammad Reza Koohkan, Marc Bocquet, Lin Wu, Monika Krysta [The Preparatory Commission for the Comprehensive Nuclear Test-Ban Treaty Organization].
The International Monitoring System (IMS) radionuclide network enforces the Comprehensive Nuclear-Test-Ban Treaty, which bans nuclear explosions. We have evaluated the potential of the IMS radionuclide network for inverse modeling of the source, whereas it is usually assessed by its detection capability. To do so, we have chosen the Degrees of Freedom for the Signal (DFS), a well established criterion in remote sensing, in order to assess the performance of an inverse modeling system. Using a recent multiscale data assimilation technique, we have computed optimal adaptive grids of the source parameter space by maximizing the DFS. This optimization takes into account the monitoring network, the meteorology over one year (2009) and the relationship between the source parameters and the observations derived from the FLEXPART Lagrangian transport model. Areas of the domain where the grid-cells of the optimal adaptive grid are large emphasize zones where the retrieval is more uncertain, whereas areas where the grid-cells are smaller and denser stress regions where more source variables can be resolved.

The observability of the globe through inverse modeling is studied in strong, realistic and small model error cases. The strong error and realistic error cases yield heterogeneous adaptive grids, indicating that information does not propagate far from the monitoring stations, whereas in the small error case, the grid is much more homogeneous. In all cases, several specific continental regions remain poorly observed such as Africa as well as the tropics, because of the trade winds. The northern hemisphere is better observed through inverse modeling (more than 60% of the total DFS) mostly because it contains more IMS stations. This unbalance leads to a better performance of inverse modeling in the northern hemisphere winter. The methodology is also applied to the subnetwork composed of the stations of the IMS network that measure noble gases.

### 6.3.3. Optimal redistribution of the background ozone monitoring stations over France

**Participants:** Lin Wu, Marc Bocquet.

In this study, the BDQA background stations are partially redistributed over France under a set of design objectives, which are defined on a regular grid that covers France. Spatial interpolations are used to extrapolate simulated concentrations (of chemistry-transport models or assimilation results) to these grid nodes. Three types of criteria are considered: the geostatistical, geometrical, and physical ones. Simulated annealing is employed to select optimally the stations. Significant improvement with all the proposed criteria has been found for the optimally redistributed network against the original background BDQA network. For complex objectives, e.g. that addressing the heterogeneity of ozone field, the physical criteria are more appropriate.

### 6.4. Ensemble methods

Due to the great uncertainties that arise in air quality modeling, relying on a single model may be not sufficient. Therefore ensembles of simulations are now considered in a wide range of applications, from uncertainty estimation to operational forecast.

#### 6.4.1. Uncertainty estimation based on multimodel ensembles and Monte Carlo simulations

**Participants:** Damien Garaud, Vivien Mallet, Raphaël Périllat, Irène Korsakissok [IRSN], Denis Quélo [IRSN].

Air quality forecasts are limited by strong uncertainties especially in the input data and in the physical formulation of the models. There is a need to estimate these uncertainties for the evaluation of the forecasts, the production of probabilistic forecasts, and a more accurate estimation of the error covariance matrices required by data assimilation.

Because a large part of the uncertainty in the forecast originates from uncertainties in the model formulation (primarily the physical parameterizations), a multimodel ensemble seems to be the adequate tool for uncertainty estimation. Several 100-member ensembles were generated and calibrated (based on observations – see the example in Figure 4) over year 2007 in order to study the impact of EDF thermal plants on air quality. The ensemble simulations were carried out at European scale and at the scale of French regions so as to study the local impact of two EDF plants. This approach allowed us to estimate the uncertainty on the simulated impact of the plants. Based on the calibrated ensembles, we also computed probabilistic forecasts for threshold exceedences.
Specific work has been carried out to decompose the discrepancy between observations and model simulations in three errors: the modeling error, the representativeness error and the observational error. It was shown that the representativeness error might account for a third of the discrepancy.

After the Fukushima nuclear disaster, we worked on uncertainty estimation of the transport simulations at Japan scale. We carried out Monte Carlo simulations with perturbations on most input parameters. The first results showed that the variability of the ensemble simulations is great enough to reasonably sample the uncertainties after a calibration.

![Figure 4. Uncertainty estimate (standard deviation) of two ensembles: a raw ensemble on the left, a calibrated ensemble on the right. The uncertainty estimation is significantly reduced after calibration.](image-url)

### 6.4.2. Aggregation of meteorological forecasts

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

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). The generation of a single forecast with a weighted linear combination is called sequential aggregation. Each time new observations are available, the weights of the linear combination are updated and applied for the next forecasts. We applied the discounted ridge regression algorithm, which we previously introduced for sequential aggregation of air quality forecasts, to forecast wind and temperature at given observation stations. The ensemble was generated with forecasts at different range from two models. The aggregation proved to be efficient for one-day forecasts at least.

### 6.5. Image assimilation

Sequences of images, such as satellite acquisitions, display structures evolving in time. This information is recognized of major interest by forecasters (Meteorologists, oceanographers, etc) in order to improve the information provided by numerical models. However, these satellite images are mostly assimilated in geophysical models on a point-wise basis, discarding the space-time coherence visualized by the evolution of structures such as clouds. Assimilating in an optimal way image data is of major interest and this issue should be considered in two ways:
• from the model’s viewpoint, the problem is to control the location of structures using the observations,
• from the image’s viewpoint, a model of the dynamics and structures has to be built from the observations.

6.5.1. Divergence-free motion estimation

Participants: Dominique Béréziat [UPMC/LIP6], Isabelle Herlin, Nicolas Mercier, Sergiy Zhuk.

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

Motion is characterized by its vorticity value and assumed to satisfy the Lagragian constancy hypothesis. An image model is then defined: the state vector includes the vorticity, whose evolution equation is derived from that of motion, and a pseudo-image that is transported by motion. An image assimilation method, based on the 4D-Var technics, is defined and developed that estimates motion as a compromise between the evolution equations of vorticity and pseudo-image and the observed sequence of images: the pseudo-images have to be similar to the acquisitions.

As the evolution equations of vorticity and pseudo-image involve the motion value, the motion field has to be retrieved at each time step of the studied temporal window. An algebraic method, based on the projection of vorticity on a subspace of eigenvectors of the Laplace operator, is defined in order to allow Dirichlet boundary conditions for the vorticity field.

The divergence-free motion estimation method is tested and quantified on synthetic data. This shows that it computes a quasi-exact solution and outperforms the state-of-the-art methods that were applied on the same data.

The method is also applied on Sea Surface Temperature (SST) images acquired over Black Sea by NOAA-AVHRR sensors. The divergence-free assumption is roughly valid on these acquisitions, due to the small values of vertical velocity at the surface. Fig. 5 displays data and results. As no ground truth of motion is available, the method is quantified by the value of correlation between the pseudo-images and the real acquisitions. Again, the method provides the best result compared to other state-of-the-art algorithms.

6.5.2. Improvement of motion estimation by assessing errors on the dynamics

Participants: Dominique Béréziat [UPMC/LIP6], Isabelle Herlin, Nicolas Mercier.

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, we consider adding a model error term in the evolution equation of motion and design a weak formulation of 4D-Var data assimilation. The cost function to be minimized simultaneously depends 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.

This error assessment method is evaluated and quantified on twin experiments, as no ground truth would be available for real image data. Fig. 6 shows four frames of a series of observations obtained by integrating the evolution model from an initial condition on image and velocity field (the ground truth \( w_{\text{ref}}(0) \) displayed on the left of Fig. 7). An error value is added at each time step on the motion value, when integrating the simulation model. This error is a constant bias.

We performed two data assimilation experiments. The first one considers the evolution model as perfect, with no error in the evolution equation. It is denoted PM (for Perfect Model). The second one, denoted IM (for Imperfect Model) involves an error in the motion evolution equation. In fig. 7 are displayed the motion fields retrieved by PM and IM at the beginning of the temporal window.

As it can be seen, IM computes a correct velocity field while PM completely fails.
Figure 5. Results on SST images (observation and motion result).

Figure 6. Observations Images.

Figure 7. Comparaison estimations with ground truth.
The results on this error assessment method are still preliminary. Perspectives are considered in order to correctly retrieve the error on dynamics by constraining its shape. An important application is, for instance, the detection of dynamics changes on long temporal sequences.

6.5.3. Nonlinear Observation Equation For Motion Estimation

Participants: Dominique Béréziat [UPMC/LIP6], Isabelle Herlin.

In the image processing literature, the optical flow equation is usually chosen to assess motion from an image sequence. However, it corresponds to an approximation that is no more valid in case of large displacements. We evaluate the improvements obtained when using the non linear transport equation of the image brightness by the velocity field. A 4D-Var data assimilation method is designed that simultaneously solves the evolution equation and the observation equation, in its non linear and linearized form. The comparison of results obtained with both observation equations is quantified on synthetic data and discussed on oceanographic Sea Surface Temperature (SST) images. We show that the non linear model outperforms the linear one, which underestimates the motion norm. Fig. 8 illustrates this on SST images (motion vectors are displayed by arrows).

![Figure 8. Results on SST images. Top: Non-linear observation equation. Bottom: linear.](image)

The aim of this research is to achieve a correct estimation of motion when the object displacement is greater than its size. In this case, coarse-to-fine incremental methods as well as the non linear data assimilation method fail to retrieve a correct value. The perspective is then to include, in the state vector, a variable describing the trajectory of pixels. The observation operator will then measure the effective displacement of pixels, according to their trajectories, and allow a better estimation of motion value.

6.5.4. Recovering missing data on images

Participants: Dominique Béréziat [UPMC/LIP6], Isabelle Herlin, Nicolas Mercier.
A data assimilation method was designed to recover missing data and reduce noise on satellite acquisition. The state vector includes motion and image fields. Its evolution equation is based on assumptions on the underlying dynamics displayed by the sequence of images and considers the passive transport of images by the velocity field. The observation equation compares the image component of the state vector and the real observations. Missing and noisy data regions are characterized by a gaussian observation error, whose covariance matrix $R$ has an approximately infinitesimal inverse. The noise recovering method computes a solution of the state vector that is a compromise between the evolution equation and the observation equation. The image component of the solution satisfies the assumptions on the dynamics and is close to the real acquisition according to the covariance matrix $R$. This image component provides the reconstruction of the noisy acquisitions.

The recovering method was applied on synthetically noised SST images in order to quantify the quality of the recovering (see Fig. 9).

![Figure 9. Recovering of noisy data.](image)

The method is a promising alternative to those such as space-time interpolation. In the experiments, the Lagrangian constancy of the state vector is used as evolution equation. The perspectives concern the use of more advanced dynamic equations, as for instance the shallow water equations that link the motion field to the thickness of the ocean surface layer, and improved modeling of illumination changes over the sequence, due to various order acquisition times.

6.5.5. Validation of velocity estimated with image assimilation

**Participants:** Isabelle Herlin, Etienne Huot, Gennady Korotaev [Marine Hydrophysical Institute, Ukraine], Evgeny Plotnikov [Marine Hydrophysical Institute, Ukraine].

This study is achieved in collaboration with the Marine Hydrophysical Institute (MHI) of Sevastopol. The aim is to estimate and further validate the estimation of Black Sea surface velocity from sequences of satellite images in order to allow an optimal assimilation of these pseudo-observations in 3D ocean circulation models. Several Image Models were designed that express the dynamics of velocity and the temporal evolution of image data. An image assimilation method was developed based on the 4D-Var formalism and estimates motion as a compromise between the Image Model, the image acquisitions and regularity heuristics on the velocity field. Two Image Models were qualitatively and quantitatively compared: the *Stationary Image Model* (SIM) based on the heuristics of stationary motion, which is valid at short temporal scale, and the *Shallow Water Image Model* (SWIM), based on the shallow-water equations.

The comparison between SIM and SWIM results confirms that SIM provides correct results only on short temporal windows, while SWIM allows to process longer image sequences.
The validation of motion estimation by image assimilation requires additional observation data, as no measure of motion is available from satellite sensors. Sea Level Anomaly, measured by satellite altimeters, is then compared to the thickness of the surface layer as estimated by the Shallow Water Image Model. This comparison shows a good adequacy of shape and values [30], [32]. As the velocity field is strongly related to this thickness value from the physical evolution laws, these results further validate the estimation of the velocity and the image assimilation approach.

### 6.5.6. Velocity estimation under the geostrophic equilibrium assumption

**Participants:** Isabelle Herlin, Etienne Huot.

The surface motion of the Black Sea approximately verifies the geostrophic equilibrium property. As the surface velocity can be directly derived from the surface layer thickness $h$, this allows to simplify the shallow-water equations and the dynamics is expressed by the evolution of $h$. The Geostrophic Shallow Water Image Model (GSWIM) is then designed based on the evolution of $h$ and the image data. A 4D-Var assimilation method was designed and developed in order to estimate $h$ from a sequence of satellite images. The motion field is then computed from the estimation of $h$.

This method was first tested and quantified on twin experiments with satellite data. Figure 10 simultaneously displays the result of the velocity estimation by GSWIM and the ground truth.

![Image](image1.png)

**Figure 10.** Left: first image of the sequence, Center: motion estimated by GSWIM, Right: ground truth.

### 6.5.7. Coupling models for motion estimation on long temporal image sequences

**Participants:** Karim Drifi, Isabelle Herlin.

This study concerns the estimation of motion fields from satellite images on long temporal sequences. The huge computational cost and memory required by data assimilation methods on the pixel grid makes impossible to use these techniques on long temporal intervals. For a given dynamic model (named full model on the pixel grid), the Galerkin projection on a subspace provides a reduced model that allows image assimilation at low cost. The definition of this reduced model however requires defining the optimal subspace of motion. A sliding windows method is thus designed:

- The long image sequence is split into small temporal windows that half overlap in time.
- Data assimilation in the full model is applied on the first window to retrieve the motion field.
- The estimate of motion field at the beginning of the second window makes it possible to define the subspace for motion and a reduced model is obtained by Galerkin projection.
- Data assimilation in the reduced model is applied for this second window.
- The process is then iterated for the next window until the end of the whole image sequence.
Figure 11 summarize the described methodology.

Twin experiments were designed to quantify the results of this sliding windows method. Results on motion estimation are given in Figure 12 and compared with the ground truth. The NRMSE (in percentage) ranges from 1.1 to 4.0% from the first to the sixth window. On the first window, 3 hours are required to estimate the motion fields with the full model. For the next 5 windows, less than 1 minute is required to compute motion.
6.6. Minimax filtering

Participants: Vivien Mallet, Sergiy Zhuk.

In air quality modeling, the model error is supposed to take into account the uncertainty on the meteorological fields (winds and vertical diffusivities), the segregation and mixing in emission areas that affect the effective kinetic rates of reactions, the boundary condition fields, all physical parameterizations (dry deposition, wet scavenging), etc. All the above sources of error have bounded energy and typically are not normally distributed or independent.

In order to take this into account in the data assimilation process, we applied the Minimax State Estimation (MSE) approach. It is well known that a bottle-neck of minimax estimation algorithms as well as of the family of Kalman-type filters is the dimension issue. To solve it, we applied a powerful version of the minimax filter developed for the so-called differential-algebraic equations. This filter works for any linear ordinary differential equation with time-dependent coefficients on any linear manifold, which can also change in time. Based on this novel approach, we derived a computationally tractable reduced version of the minimax filter. The derivation was made in a new and rigorous framework. In addition to the reduction, the new filter shows all the interesting properties inherited from the minimax setting, especially the description of the model and observational errors, which only need to have bounded energy. The later is important in the context of applications because the errors are always bounded. In contrast, most high-dimensional statistical filters are designed for unbounded random errors with special distribution function.

The algorithm, already implemented in the data assimilation library Verdandi, was further developed to compute a better reduction base.

The algorithm was in addition applied for ensemble sequential aggregation. The minimax filter computes weights for each model in the ensemble and a forecast is generated as the weighted linear combination of the ensemble members. In this case, the dimension is small so that no reduction is needed. The approach shows two noteworthy advantages: the observational errors can be taken into account and a dynamics can be given for the weights.

6.6.1. A posteriori minimax motion estimation

Participants: Sergiy Zhuk, Isabelle Herlin.

Data assimilation algorithms based on the 4D-Var formulation look for the so-called conditional mode estimate. The latter maximizes the conditional probability density function, provided the initial condition, model error and observation noise are realizations of independent Gaussian random variables. However this Gaussian assumption is often not satisfied for geophysical flows. Moreover, the estimation error of the conditional mode estimate is not a first-hand result of these methods. The issues above can be addressed by means of the Minimax State Estimation (MSE) approach. It allows to filter out any random (with bounded correlation operator) or deterministic (with bounded energy) noise and assess the worst-case estimation error.

The iterative MSE algorithm was developed for the problem of optical flow estimation from a sequence of 2D images. The main idea of the algorithm is to use the "bi-linear" structure of the Navier-Stokes equations and optical flow constraint in order to iteratively estimate the optical flow. The algorithm consists of the following parts:

1) we construct the pseudo-observations that is the estimate of the image brightness function \( \hat{I}(x, y, t) \) solving the optical flow constraint such that \( \hat{I} \) fits (in the sense of least-squares) the observed sequence of images; to do so we set the velocity field in the optical flow constraint to be the current minimax estimate of the velocity field \( \hat{v} \), obtained at the previous iteration of the algorithm, and construct the minimax estimate \( \hat{I} \) of the solution of the resulting linear advection equation using image sequence as discrete measurements;

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

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

Currently numerical experiments are carried out in order to study the convergence rate of the algorithm.

6.7. Fire application

6.7.1. Model evaluation for fire propagation

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

In the field of forest fires risk management, important challenges exist in terms of people and goods preservation. Answering to strong needs from different actors (firefighters, foresters), researchers focus their efforts to develop operational decision support system tools that may forecast wildfire behavior. This requires the evaluation of models performance, but currently, simulation errors are not sufficiently qualified and quantified. As the main objective is to realize a decision support system, it is required to establish robust forecast evaluations. In the context of the ANR project IDEA, the evaluation of model simulations has been started with a bibliographical review, the implementation of a series of forecast scores and the definition of a series of ideal cases where some classical scores may fail (especially in taking into account the dynamics).

In addition, we consider that the proper evaluation of a model requires to apply it to a large number of fires – instead of carrying out a fine tuning on just one fire. We implemented a software to simulate a large number of fires (from the Prométhée database, http://www.promethee.com/ ) with the simulation model ForeFire (CNRS/University of Corsica) and to evaluate the results with our error measures. One simulation requires mainly the following data: the ignition point, the ground elevation, the vegetation cover and the wind field. See the illustration of Fig. 13.

![Fire simulation](image)

Figure 13. Fire simulation (using ForeFire) in red elevated contour, and observation (from Prométhée) of the burned area in filled red contour, for a 2003 fire near San-Giovanni-di-Moriani (Corsica).
ESTIME Project-Team (section vide)
6. New Results

6.1. Modeling and estimation in biomechanics

6.1.1. Patient-Specific Electromechanical Models of the Heart for the Prediction of Pacing Acute Effects in CRT: a Preliminary Clinical Validation

In collaboration with Project-Team Asclepios from INRIA Sophia-Antipolis-Méditerranée and the Division of Imaging Sciences of St Thomas’ Hospital, King’s College London we demonstrated the benefits of using patient-specific electromechanical models of the heart for the prediction of pacing acute effects in CRT, see [5].

Cardiac resynchronisation therapy (CRT) is an effective treatment for patients with congestive heart failure and a wide QRS complex. However, up to 30% of patients are non-responders to therapy in terms of exercise capacity or left ventricular reverse remodelling. A number of controversies still remain surrounding patient selection, targeted lead implantation and optimisation of this important treatment. The development of biophysical models to predict the response to CRT represents a potential strategy to address these issues. We present how the personalisation of an electromechanical model of the myocardium can predict the acute haemodynamic changes associated with CRT. In order to introduce such an approach as a clinical application, we needed to design models that can be individualised from images and electrophysiological mapping of the left ventricle. We performed the personalisation of the anatomy, the electrophysiology, the kinematics and the mechanics. The acute effects of pacing on pressure development were predicted with the in silico model for several pacing conditions on two patients, achieving good agreement with invasive haemodynamic measurements: the mean error on \( \frac{dP}{dt_{\max}} \) is \( 47.5 \pm 35 \) mmHg.s\(^{-1}\), less than 5% error.

6.1.2. Estimation of tissue contractility from cardiac MRI using a biomechanical heart model

Participants: Radomir Chabiniok, Dominique Chapelle, Alexandre Imperiale, Philippe Moireau.

In collaboration with P.-F. Lesault, A. Rahmouni and J.-F. Deux from Hospital H. Mondor, Créteil we proposed and assessed an estimation procedure – based on data assimilation principles – well-suited to obtain some regional values of key biophysical parameters in a beating heart model, using actual Cine-MR images, see [8], [1]. The motivation is twofold: (1) to provide an automatic tool for personalizing the characteristics of a cardiac model in order to achieve predictivity in patient-specific modeling, and (2) to obtain some useful information for diagnosis purposes in the estimated quantities themselves. In order to assess the global methodology we specifically devised an animal experiment in which a controlled infarct was produced and data acquired before and after infarction, with an estimation of regional tissue contractility – a key parameter directly affected by the pathology – performed for every measured stage. After performing a preliminary assessment of our proposed methodology using synthetic data, we then demonstrate a full-scale application by first estimating contractility values associated with 6 regions based on the AHA subdivision, before running a more detailed estimation using the actual AHA segments. The estimation results are assessed by comparison with the medical knowledge of the specific infarct, and with late enhancement MR images. We discuss their accuracy at the various subdivision levels, in the light of the inherent modeling limitations and of the intrinsic information contents featured in the data.

We are now working on improving these results by the use of Tagged-MRI, see [9]. The first approach consists in assuming that the image data is processed in the form of deforming tag planes, which we employ to obtain a discrepancy between the model and the data by computing distances to these surfaces. We assess our procedure using synthetic measurements produced with a model representing an infarcted heart as observed in an animal experiment, and the estimation results are found to be of superior accuracy compared to assimilation based on segmented endo- and epicardium surfaces. Then we extend this strategy to tagged lines instead of tagged planes or even directly with apparent displacements extracted from tagged images by optical flow methods.
6.1.3. Convergence of observers based on partial field measurements for the wave equation

Participants: Dominique Chapelle, Nicolae Cîndea, Maya de Buhan, Philippe Moireau.

We analyzed an observer strategy based on partial – i.e. in a subdomain – measurements of the solution of a wave equation, in order to compensate for unknown initial conditions, see \[ 17 \], \[ 18 \]. We proved the exponential convergence of this observer under a non-standard observability condition, whereas using measurements of the time-derivative of the solution would lead to a standard observability condition arising in stabilization and exact controllability. Nevertheless, we directly related our specific condition to the classical geometric control condition. This results justify in a linear framework the use of our observer-based filter in cardiac modeling.

6.1.4. Reduced nonlinear optimal filtering

Participants: Dominique Chapelle, Akos Matszangosz, Philippe Moireau.

We investigated some issues pertaining to reduced-order considerations in nonlinear optimal filtering. Classically, optimal filtering formulations lead to Hamilton-Jacobi-Bellman (HJB) equations posed in the complete “space of uncertainty”, namely, including the state space. This makes such methods generally untractable for PDE-based models. However, under certain assumptions pertaining to reduced uncertainties we can transform the HJB equations into a form posed in the reduced uncertainty space, and with only time derivatives involved. This form can be solved for – including with PDEs – provided this reduced space is of limited size, and then gives a reference “optimal” method to which other filtering procedures can be compared. The subject of Akos Matszangosz’ internship (from “Ecole des Mines de Paris”, duration 4 months) was to design and perform an adequate implementation of this reduced-order optimal filter, based on a sparse-grid discretization of the uncertainty space. In addition, we are currently working on discrete-time optimal filtering formulations, which are distinct – and preferable in principle – to discretizing the continuous forms.

6.2. Asymptotic and multiscale modeling

6.2.1. Modeling and simulation of multi-layers mechanical structures

Participants: Marina Vidrascu, Sofiene Hendili.

The collaboration with Françoise Krasucki (Montpellier University) and Giuseppe Geymonat (Ecole polytechnique) on the modeling of 3D materials connected by stiff interfaces continues within the Epsilon ANR project (Domain decomposition and multi-scale computations of singularities in mechanical structures 7.1.1 ). In the framework of matched asymptotic expansions we introduced a new effective and robust method to approximate the behavior of a structure containing a thin layer with periodically distributed heterogeneities (holes, rigid bodies...), see \[ 3 \], \[ 10 \], \[ 11 \].

6.2.2. Multi-scale modeling and simulation of rubber

Participants: Maya de Buhan, Marina Vidrascu, Antoine Gloria [SIMPAF], François Lequeux [ESPCI], Patrick Le Tallec [Ecole Polytechnique].

In collaboration with A. Gloria (project-team SIMPAF) and P. Le Tallec (Ecole Polytechnique), we are currently working on a multiscale model for rubber based on the statistical physics description of a network of polymer chains. The numerical simulation of the model has been addressed within the ARC Disco using the Sheldon software. Comparisons with mechanical experiments are promising, and related inverse problems have been addressed in the post-doc of M. de Buhan. Two publications are in preparation.

This work is supported by the ARC DISCO (7.1.2).

6.3. Other topics

6.3.1. Numerical analysis of POD-based Galerkin approximations

Participants: Dominique Chapelle, Asven Gariah, Philippe Moireau, Jacques Sainte-Marie.
In [2], we proposed a numerical analysis of Proper Orthogonal Decomposition (POD) model reductions in which a priori error estimates are expressed in terms of the projection errors that are controlled in the construction of POD bases. These error estimates are derived for generic parabolic evolution PDEs, including with non-linear Lipschitz right-hand sides, and for wave-like equations. A specific projection continuity norm appears in the estimates and — whereas a general uniform continuity bound seems out of reach — we prove that such a bound holds in a variety of Galerkin bases choices. Furthermore, we directly numerically assess this bound — and the effectiveness of the POD approach altogether — for test problems of the type considered in the numerical analysis, and also for more complex equations. Namely, the numerical assessment includes a parabolic equation with super-linear reaction terms, inspired from the FitzHugh-Nagumo electrophysiology model, and a 3D biomechanical heart model. This shows that the effectiveness established for the simpler models is also achieved in the reduced-order simulation of these highly complex systems.

This work is now being continued in order to handle parameter-dependent models, and thence estimation problems.

6.3.2. Sail modeling

**Participants:** Dominique Chapelle, Daniele Trimarchi, Marina Vidrascu.

A dynamic Finite Element method – based on non-linear MITC shell finite elements implemented in the MITCNL software – has been proposed and assessed for the analysis of downwind sail-type structures, see [12]. The main purpose was here to investigate the development of wrinkling, a phenomenon commonly observed in practice for such structures. Considering the wrinkling in this type of analysis is of great interest, since wrinkling affects the stress distribution in the fabric. Further developments primarily regard various refinements of the model, in order to represent some even more realistic sail configurations such as with non-isotropic material models, corner reinforced zones, and cable boundary conditions. Of course, another very important perspective – and work in progress, indeed – concerns the use of such sail models coupled with the wind flow in a fluid-structure interaction framework.
5. New Results

5.1. Statistical analysis of spike trains

Modern advances in neurophysiology techniques, such as two-photon imaging of calcium signals or micro-electrode arrays (MEA) electro-physiology, have made it possible to observe simultaneously the activity of large assemblies of neurons. Such experimental recordings provide a great opportunity to unravel the underlying interactions of neural assemblies and to understand how neural populations dynamically encode information. The goal of the present project is to propose to the neuroscientists community statistical methods and numerical tools to analysing the statistics of action potentials (spike trains) obtained from MEA recordings. Our work is grounded on one hand on theoretical results on Gibbs distributions in neural networks and the other hand on a C/C++ library of algorithms developed jointly with the CORTEX INRIA team, freely available at http://enas.gforge.inria.fr/. We have collaborations with several labs specialized in MEA recording from the retina: Centro Interdisciplinario de Neurociencia de Valparaiso, Universidad de Valparaiso, Chile http://www.cinv.cl/; Department of Molecular Biology and Princeton Neuroscience Institute, Princeton University, USA http://www.princeton.edu/neuroscience/; Institut de la Vision, Paris http://www.institut-vision.org/.

5.1.1. A discrete time neural network model with spiking neurons. Dynamics with noise.

Participant: Bruno Cessac [correspondent].

We provide rigorous and exact results characterizing the statistics of spike trains in a network of leaky Integrate-and-Fire neurons, where time is discrete and where neurons are submitted to noise, without restriction on the synaptic weights. We show the existence and uniqueness of an invariant measure of Gibbs type and discuss its properties. We also discuss Markovian approximations and relate them to the approaches currently used in computational neuroscience to analyse experimental spike trains statistics. This work has appeared in Journal of Mathematical Biology[17].

5.1.2. Statistics of spike trains in conductance-based neural networks: Rigorous results

Participant: Bruno Cessac [correspondent].

We consider a conductance-based neural network inspired by the generalized Integrate and Fire model introduced by Rudolph and Destexhe in 1996. We show the existence and uniqueness of a unique Gibbs distribution characterizing spike train statistics. The corresponding Gibbs potential is explicitly computed. These results hold in the presence of a time-dependent stimulus and apply therefore to non-stationary dynamics. This establishes a rigorous ground for the current investigations attempting to characterize real spike trains data with Gibbs distributions, such as the Ising-like distribution, using the maximal entropy principle. This work has appeared in Journal of Mathematical Neuroscience[18].


Participants: Bruno Cessac [correspondent], Adrian Palacios [Centro de Neurociencia, Valparaiso, Chile].

This work focuses on methods from statistical physics and probability theory allowing the analysis of spike trains in neural networks. Taking as an example the retina we present recent works attempting to understand how retina ganglion cells encode the information transmitted to the visual cortex via the optical nerve, by analyzing their spike train statistics. We compare the maximal entropy models used in the literature of retina spike train analysis to rigorous results establishing the exact form of spike train statistics in conductance-based Integrate-and-Fire neural networks. This work is submitted in “Mathematical Problems in Computational Biology and Biomedicine” Springer, [54].
5.1.4. Gibbs distribution analysis of temporal correlations structure in retina ganglion cells

**Participants:** Michael Berry II [Department of Molecular Biology, Princeton University, USA], Bruno Cessac [correspondent], Olivier Marre, Adrian Palacios [Centro de Neurociencia, Valparaiso, Chile], Juan-Carlos Vasquez.

We present a method to estimate Gibbs distributions with spatio-temporal constraints on spike trains statistics. We apply this method to spike trains recorded from ganglion cells of the salamander retina, in response to natural movies. Our analysis, restricted to a few neurons, performs more accurately than pairwise synchronization models (Ising) or the 1-time step Markov models (Marre et al. (2009)) to describe the statistics of spatio-temporal spike patterns and emphasizes the role of higher order spatio-temporal interactions. This work has been accepted in *Journal of Physiology, Paris* [28] (in press).

5.1.5. A Markovian event-based framework for stochastic spiking neural networks

**Participants:** Olivier Faugeras, Jonathan Touboul.

In spiking neural networks, the information is conveyed by the spike times, that depend on the intrinsic dynamics of each neuron, the input they receive and on the connections between neurons. In this article we study the Markovian nature of the sequence of spike times in stochastic neural networks, and in particular the ability to deduce from a spike train the next spike time, and therefore produce a description of the network activity only based on the spike times regardless of the membrane potential process. To study this question in a rigorous manner, we introduce and study an event-based description of networks of noisy integrate-and-fire neurons, i.e. that is based on the computation of the membrane potential process. We show that the firing times of the neurons in the networks constitute a Markov chain whose transition probability is related to the probability distribution of the interspike interval of the neurons in the network. In the cases where the Markovian model can be developed, the transition probability is explicitly derived in such classical cases of neural networks as the linear integrate-and-fire neuron models with excitatory and inhibitory interactions, for different types of synapses, possibly featuring noisy synaptic integration, transmission delays and absolute and relative refractory period. This covers most of the cases that have been investigated in the event-based description of spiking deterministic neural networks.

This work has appeared in *Journal of Computational Neuroscience* [26].

5.2. Coding by spikes

Our goal here is a better understanding of the extent to which computing and modeling with spiking neuron networks might be biologically plausible and computationally efficient. Based on a thorough characterization of the main constraints on spiking neural networks dynamics this has led us to propose new algorithms to infer the structure of the network from its spike trains and to propose an FPGA implementation of spiking neural networks.

5.2.1. Reverse-engineering of spiking neural networks parameters

**Participants:** Bruno Cessac [correspondent], Horacio Rostro-Gonzalez, Thierry Viéville [Cortex].

We consider the deterministic evolution of a time-discretized spiking network of neurons with connection weights having delays, modeled as a discretized neural network of the generalized integrate and fire (gIF) type. The purpose is to study a class of algorithmic methods allowing to calculate the proper parameters (synaptic weights) to reproduce exactly a given spike train generated by an hidden (unknown) neural network. This problem is linear (L) if the membrane potentials are observed and LP (Linear-Programming) if only spike times are observed, in the context of gIF models. The L or LP adjustment mechanism is local to each unit and has the same structure as an "Hebbian" rule. This paradigm is easily generalizable to the design of input-output spike train transformations. This means that we have a practical method to "program" a spiking network, i.e. find a set of parameters allowing us to exactly reproduce the network output, given an input.

This work has been submitted in the *Journal of Neural Engineering*, 2011[25].
5.2.2. Development of FPGA-based efficient reconfigurable architectures for spiking neural networks

Participants: Bruno Cessac, Bernard Girau [INRIA Cortex], Horacio Rostro-Gonzalez, Cesar Torres-Huitzil [Information Technology Department, Polytechnic University of Victoria (UPV), Tamaulipas, Mexico], Thierry Viéville [Cortex, correspondent].

Spiking neural networks are able to perform very powerful computations with precise timed spikes. We are developing an FPGA (Field Programmable Gate Array) reconfigurable platform that enables the simulation of in silico models of spiking neural networks. Since the model is directly mapped into a FPGA device, the neural processing is accelerated and the time consumption reduced. We use VHDL and Handel-C to design the reconfigurable architecture of a discrete time Integrate-and-Fire model coded in CUDA, running on GPU. This work has been accepted in Journal of Physiology, Paris [24].

5.2.3. Towards biologically inspired image coders

Participants: Marc Antonini [Laboratoire I3S, Sophia Antipolis, France], Pierre Kornprobst, Khaled Masmoudi [Laboratoire I3S, Sophia Antipolis, France].

In [51] we presented a novel bio-inspired and dynamic coding scheme for static images. Our coder aims at reproducing the main steps of the visual stimulus processing in the mammalian retina taking into account its time behavior. The main novelty of this work is to show how to exploit the time behavior of the retina cells to ensure, in a simple way, scalability and bit allocation. To do so, our main source of inspiration has been the biologically plausible retina model Virtual Retina described in Section 4.1. Following a similar structure, our model has two stages. The first stage is an image transform which is performed by the outer layers in the retina. Here it is modelled by filtering the image with a bank of difference of Gaussians with time-delays. The second stage is a time-dependent analog-to-digital conversion which is performed by the inner layers in the retina. Thanks to its conception, our coder enables scalability and bit allocation across time. Also, compared to the JPEG standards, our decoded images do not show annoying artefacts such as ringing and block effects. As a whole, this article shows how to capture the main properties of a biological system, here the retina, in order to the design a new efficient coder.

5.3. Mean field methods

5.3.1. Noise-induced behaviors in neural mean field dynamics

Participants: Jonathan Touboul, Geoffroy Hermann, Olivier Faugeras.

The collective behavior of cortical neurons is strongly affected by the presence of noise at the level of individual cells. In order to study these phenomena in large-scale assemblies of neurons, we consider networks of firing-rate neurons with linear intrinsic dynamics and nonlinear coupling, belonging to a few types of cell populations and receiving noisy currents. Asymptotic equations as the number of neurons tends to infinity (mean field equations) are rigorously derived based on a probabilistic approach. These equations are implicit on the probability distribution of the solutions which generally makes their direct analysis difficult. However, in our case, the solutions are Gaussian, and their moments satisfy a closed system of nonlinear ordinary differential equations (ODEs), which are much easier to study than the original stochastic network equations, and the statistics of the empirical process uniformly converge towards the solutions of these ODEs. Based on this description, we analytically and numerically study the influence of noise on the collective behaviors, and compare these asymptotic regimes to simulations of the network. We observe that the mean field equations provide an accurate description of the solutions of the network equations for network sizes as small as a few hundreds of neurons. In particular, we observe that the level of noise in the system qualitatively modifies its collective behavior, producing for instance synchronized oscillations of the whole network, desynchronization of oscillating regimes, and stabilization or destabilization of stationary solutions. These results shed a new light on the role of noise in shaping collective dynamics of neurons, and gives us clues for understanding similar phenomena observed in biological networks.
This work has been accepted for publication in the SIAM Journal on Applied dynamical Systems [72].

5.3.2. Mean Field description of and propagation of chaos in recurrent multipopulation networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons

Participants: Javier Baladron, Diego Fasoli, Olivier Faugeras, Jonathan Touboul.

We derive the mean-field equations arising as the limit of a network of interacting spiking neurons, as the number of neurons goes to infinity. The neurons belong to a fixed number of populations and are represented either by the Hodgkin-Huxley model or by one of its simplified version, the Fitzhugh-Nagumo model. The synapses between neurons are either electrical or chemical. The network is assumed to be fully connected. The maximum conductances vary randomly. Under the condition that all neurons initial conditions are drawn independently from the same law that depends only on the population they belong to, we prove that a propagation of chaos phenomenon takes place, namely that in the mean-field limit, any finite number of neurons become independent and, within each population, have the same probability distribution. This probability distribution is solution of a set of implicit equations, either nonlinear stochastic differential equations resembling the McKean-Vlasov equations, or non-local partial differential equations resembling the McKean-Vlasov-Fokker-Planck equations. We prove the well-posedness of these equations, i.e. the existence and uniqueness of a solution. We also show the results of some preliminary numerical experiments that indicate that the mean-field equations are a good representation of the mean activity of an infinite size network, even for modest sizes. These experiment also indicate that the McKean-Vlasov-Fokker-Planck equations may be a good way to understand the mean-field dynamics through, e.g., a bifurcation analysis.

This work has been submitted for publication in the Journal of Mathematical Neuroscience [55].

5.3.3. Three applications of GPU computing in neuroscience

Participants: Javier Baladron, Olivier Faugeras.

GPUs are low cost highly parallel devices that are now not only used for graphics but also for numerical simulation. We present three applications of a computer system with multiple GPUs to the domain of theoretical neuroscience. The first application is to a continuous model of the primary visual area, the second to the simulation of a stochastic neural network, and the third to the computation of the probability distribution on the possible states of a network. In all three cases we show that the speed-up obtained by the use of GPUs has considerably helped answering a scientific or technological question.

This work has been accepted for publication in Computing in Science and Engineering [63].

5.4. Neural Fields

5.4.1. Modelling the dynamics of contextual motion integration in the primate

Participants: Heiko Neumann [Institute of Neural Information Processing, Ulm University, Ulm, Germany], Pierre Kornprobst, Guillaume Masson [Institut de Neurosciences de la Timone, UMR 6193, CNRS, Marseille, France], Emilien Tlapale.

The dynamics of motion integration show striking similarities when observed at neuronal, psychophysical, and oculomotor levels. Based on the inter-relation and complementary insights given by those dynamics, our goal is to investigate how basic mechanisms of dynamical cortical processing can be incorporated in a dynamical model to solve several aspects of 2D motion integration and segmentation.

Thanks to Emilien Tlapale PhD [13] (see also [16]), we have obtained the following results:

- We proposed a recurrent model of motion integration. Proposing a simple readout mechanism, we reproduced not only motion perception but also the dynamics of smooth pursuit eye movements on various line figures and gratings viewed through different apertures. Our model can also solve various contextual problems where extrinsic junctions should be eliminated, without relying on complex junction detectors or depth computation [71]. Finally, we have also shown how our model can be rewritten in the neural fields formalism (see [52] and the Software MotionLib), which has opened new perspectives as detailed in Section 5.4.2.
- We confronted our results to artificial and biological vision. To formalize the comparison against visual performance, we proposed a new evaluation methodology based on human visual performance by establishing a database of image sequences taken from biology and psychophysics literature \[70\], \[69\], \[67\]. We compared our results against the state of the art computer vision approaches and we found that our model also gives results comparable to recent computer vision approaches of motion estimation.

5.4.2. Neural fields models for motion integration: Characterising the dynamics of multi-stable visual motion stimuli

**Participants:** Olivier Faugeras, Pierre Kornprobst, Guillaume Masson [Institut de Neurosciences de la Timone, UMR 6193, CNRS, Marseille, France], Andrew Meso [Institut de Neurosciences de la Timone, UMR 6193, CNRS, Marseille, France], James Rankin, Emilien Tlapale, Romain Veltz.

In \[57\] we investigated the temporal dynamics of the neural processing of a multi-stable visual motion stimulus with two complementary approaches: psychophysical experiments and mathematical modelling. The so called “barber pole” stimulus is considered with an aperture configuration that supports horizontal (H), diagonal (D) or vertical (V) perceived directions for the same input. The phenomenon demonstrates an interesting variable and dynamic competition for perceptual dominance between underlying neural representations of the three directions. We probe the early processing from stimulus presentation to initial perceived direction (before perceptual reversals). Starting from a simplified neural fields model inspired from \[13\], we constructed a model of the necessary motion integration that shows a shift in perceptual dominance from D to either H or V with increasing duration. Further, the timing of this shift is shown to be controlled by a stimulus gain parameter analogous to contrast. In psychophysics experiments with concurrent eye movement recordings, observers report their perceived direction of motion for presentation durations between 0.1s and 0.5s. There is a also consistent transition in perceptual dominance from D to H/V as duration is increased. This trend, seen in both perceived direction decisions and eye movement patterns, is consistent with previous experiments using similar stimuli with an aperture configured for two (D/H) rather than three (D/H/V) states. The basic dynamic properties of the early transition from D to H/V are well predicted by the model. The experimental work additionally reveals asymmetric data patterns that guide adjustments to the model’s input equations. Observers have an H bias over V, which is also reflected in faster reaction times for H. In order to capture the bias between H and V a separate weighting is attributed to the local input corresponding to each state. The work presented forms a solid foundation for future experimental and modelling work investigating the longer term dynamics for which perceptual reversals are known to occur.

5.4.3. Analysis of a hyperbolic geometric model for visual texture perception

**Participants:** Pascal Chossat, Grégory Faye, Olivier Faugeras.

We study the neural field equations introduced by Chossat and Faugeras in \[64\] to model the representation and the processing of image edges and textures in the hypercolumns of the cortical area V1. The key entity, the structure tensor, intrinsically lives in a non-Euclidean, in effect hyperbolic, space. Its spatio-temporal behaviour is governed by nonlinear integro-differential equations defined on the Poincaré disc model of the two-dimensional hyperbolic space. Using methods from the theory of functional analysis we show the existence and uniqueness of a solution of these equations. In the case of stationary, i.e. time independent, solutions we perform a stability analysis which yields important results on their behavior. We also present an original study, based on non-Euclidean, hyperbolic, analysis, of a spatially localised bump solution in a limiting case. We illustrate our theoretical results with numerical simulations.

This work has been published in the Journal of Mathematical Neuroscience \[21\].

5.4.4. Bifurcation of Hyperbolic Planforms

**Participants:** Pascal Chossat, Grégory Faye, Olivier Faugeras.
Motivated by a model for the perception of textures by the visual cortex in primates, we analyze the bifurcation of periodic patterns for nonlinear equations describing the state of a system defined on the space of structure tensors, when these equations are further invariant with respect to the isometries of this space. We show that the problem reduces to a bifurcation problem in the hyperbolic plane \( \mathbb{D} \) (Poincaré disc). We make use of the concept of a periodic lattice in \( \mathbb{D} \) to further reduce the problem to one on a compact Riemann surface \( \mathbb{D}/\Gamma \), where \( \Gamma \) is a cocompact, torsion-free Fuchsian group. The knowledge of the symmetry group of this surface allows us to use the machinery of equivariant bifurcation theory. Solutions which generically bifurcate are called “H-planforms”, by analogy with the “planforms” introduced for pattern formation in Euclidean space. This concept is applied to the case of an octagonal periodic pattern, where we are able to classify all possible H-planforms satisfying the hypotheses of the Equivariant Branching Lemma. These patterns are, however, not straightforward to compute, even numerically, and in the last section we describe a method for computation illustrated with a selection of images of octagonal H-planforms.

This work has been published in the Journal of Nonlinear Science [19].

5.4.5. Bifurcation diagrams and heteroclinic networks of octagonal H-planforms

Participants: Grégory Faye, Pascal Chossat [correspondent].

This paper completes the classification of bifurcation diagrams for H-planforms in the Poincaré disc \( \mathbb{D} \) whose fundamental domain is a regular octagon. An H-planform is a steady solution of a PDE or integro-differential equation in \( \mathbb{D} \), which is invariant under the action of a lattice subgroup \( \Gamma \) of \( U(1, 1) \), the group of isometries of \( \mathbb{D} \). In our case \( \Gamma \) generates a tiling of \( \mathbb{D} \) with regular octagons. This problem was introduced in [19] as an example of spontaneous pattern formation in a model of image feature detection by the visual cortex where the features are assumed to be represented in the space of structure tensors. Under “generic” assumptions the bifurcation problem reduces to an ODE which is invariant by an irreducible representation of the group of automorphisms \( \mathcal{G} \) of the compact Riemann surface \( \mathbb{D}/\Gamma \). The irreducible representations of \( \mathcal{G} \) have dimension one, two, three, and four. The bifurcation diagrams for the representations of dimension less than four have already been described and correspond to already well known group actions. In the present work we compute the bifurcation diagrams for the remaining three irreducible representations of dimension four, thus completing the classification. In one of these cases, there is generic bifurcation of a heteroclinic network connecting equilibria with two different orbit types.

This work has been accepted for publication in the Journal of Nonlinear Science [22].

5.4.6. Hopf bifurcation curves in neural field networks with space-dependent delays

Participant: Romain Veltz.

We give an analytical parametrization of the curves of purely imaginary eigenvalues in the delay-parameter plane of the linearized neural field network equations with space-dependent delays. In order to determine if the rightmost eigenvalue is purely imaginary, we have to compute a finite number of such curves; the number of curves is bounded by a constant for which we give an expression. The Hopf bifurcation curve lies on these curves.

This work has appeared in the Comptes Rendus Mathématiques de l’Académie des Sciences [30].

5.4.7. Stability of the stationary solutions of neural field equations with propagation delays

Participants: Olivier Faugeras, Romain Veltz.

We consider neural field equations with space-dependent delays. Neural fields are continuous assemblies of mesoscopic models arising when modeling macroscopic parts of the brain. They are modeled by nonlinear integro-differential equations. We rigorously prove, for the first time to our knowledge, sufficient conditions for the stability of their stationary solutions. We use two methods 1) the computation of the eigenvalues of the linear operator defined by the linearized equations and 2) the formulation of the problem as a fixed point problem. The first method involves tools of functional analysis and yields a new estimate of the semigroup of the previous linear operator using the eigenvalues of its infinitesimal generator. It yields a sufficient condition for stability which is independent of the characteristics of the delays. The second method allows us to find
new sufficient conditions for the stability of stationary solutions which depend upon the values of the delays. These conditions are very easy to evaluate numerically. We illustrate the conservativeness of the bounds with a comparison with numerical simulation.

This work has appeared in the Journal of Mathematical Neuroscience [29].

5.4.8. Neural Mass Activity, Bifurcations and Epilepsy

Participants: Patrick Chauvel [INSERM U751, Marseille, Assistance Publique-Hopitaux de Marseille Timone, and Universite Aix-Marseille, Marseille], Olivier Faugeras, Jonathan Touboul, Fabrice Wendling [INSERM, U642, Rennes].

We propose a general framework for studying neural mass models defined by ordinary differential equations. By studying the bifurcations of the solutions to these equations and their sensitivity to noise we establish an important relation, similar to a dictionary, between their behaviors and normal and pathological, especially epileptic, cortical patterns of activity. We then apply this framework to the analysis of two models that feature most phenomena of interest, the Jansen and Rit model, and the slightly more complex model recently proposed by Wendling and Chauvel. This model-based approach allows to test various neurophysiological hypotheses on the origin of pathological cortical behaviors and to investigate the effect of medication. We also study the effects of the stochastic nature of the inputs which gives us clues about the origins of such important phenomena as interictal spikes, inter-ical bursts and fast onset activity, that are of particular relevance in epilepsy.

This work has appeared in Neural Computation [27].
6. New Results

6.1. Mathematical and numerical analysis of fluid-structure interaction problems

Participants: Cristóbal Bertoglio Beltran, Muriel Boulakia, Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Jimmy Mullaert.

- Over the last decade, the numerical simulation of incompressible fluid-structure interaction has been a very active research field and the subject of numerous works. In [19], we review some of the coupling schemes recently proposed in the literature. Some numerical results that show the effectiveness of the novel approaches are also presented.

- In [21], we propose a new class of time-marching schemes for the explicit coupling of an incompressible fluid and a general elastic solid (i.e., not necessarily thin [46] and possibly damped). We state a general energy-based stability result and illustrate the accuracy of the different variants in a numerical benchmark.

- [30]: This paper focuses on Eulerian-based algorithms for fluid-structure applications featuring large structural motions and/or deformations in the context of compressible flows. First, it presents a numerical method for treating simultaneously the fluid pressure and velocity conditions on static and dynamic embedded interfaces. This method is based on the exact solution of local, one-dimensional, fluid-structure Riemann problems. Next, it describes two consistent and conservative approaches for computing the flow-induced loads on rigid and flexible embedded structures.

- In [39], we present some issues encountered in fluid-structure interaction simulation (this text is targeted to a non-expert audience).

- In [42], we present a robust and efficient parameter estimation strategy for fluid-structure interaction problems. The method is based on a filtering algorithm restricted to the parameter space, known as the reduced order Unscented Kalman Filter. We illustrate our methodology with the estimation of the artery wall stiffness and the proximal Windkessel resistance.

- [44]: In this paper, we are interested in the three-dimensional coupling between an incompressible fluid and a rigid body. The fluid is modeled by the Navier-Stokes equations, while the solid satisfies the Newton’s laws. In the main result of the paper we prove that, with the help of a distributed control, we can drive the fluid and structure velocities to zero and the solid to a reference position provided that the initial velocities are small enough and the initial position of the structure is close to the reference position.

- In [46], we introduce a class of incremental displacement-correction schemes for the explicit coupling of a thin-structure with an incompressible fluid. We provide a general stability and convergence analysis that covers both the incremental and the non-incremental variants, and also the fully implicit case. The incremental variant with first-order extrapolation is unconditionally stable and yields optimal first-order accuracy in time.

6.2. Numerical methods for fluid mechanics and application to blood flows

Participants: Jean-Frédéric Gerbeau, Marc Thiriet, Irène Vignon-Clementel.
[13]: In this work, a virtual planning of three different surgical Fontan repairs was performed to test the predictive capability of a closed-loop multi-scale model (3D-0D), constructed based on preoperative patient-specific data. Results from this multi-scale approach showed that the pre-operative caval flows should not be used as boundary conditions in post-operative simulations. The Y-TCPC repair seemed to perform better than all other TCPC models both at rest and under exercise conditions. Further work is needed to correlate results from these simulations with clinical outcomes.

[18]: Flow reversal at an outlet, although perfectly physical, can lead to rapid numerical divergence in computational fluid dynamics. Several remedies have been proposed in the literature and are discussed in the present finite element study. The most robust one was found to be a boundary advective stabilization term. The comparison was done on simple examples, as well as realistic three-dimensional multi-branched models of blood flow.

[22] & [23]: Treatments for coarctation of the aorta (CoA) can alleviate blood pressure gradients, but long-term morbidity still exists that can be explained by altered indices of hemodynamics and biomechanics. These articles present a combination of CFD methods (physiologically realistic boundary conditions and FSI with viscoelastic tissue support) to explore these indices in untreated and treated CoA, comparing them to normal subjects under rest and exercise conditions. These studies showed in particular that CoA disturbs normal patterns of wall shear stress and oscillatory shear index throughout the thoracic aorta (potentially linked to the development of atherosclerosis) and that restoring favorable anatomy may not restore normal hemodynamics.

[24]: The objective of this work is to address the formulation of an adequate model of the external tissue environment when studying a portion of the arterial tree with fluid-structure interaction. The simulations are quantitatively assessed by detailed comparisons with dynamical medical image sequences, and the model results are shown to be in very good adequacy with the data.

[26]: The wide range of existing viscoelastic wall models may produce significantly different blood flow, pressure, and vessel deformation solutions in cardiovascular simulations. In this paper, we have successfully implemented and verified two viscoelastic wall models in a nonlinear 1D finite element blood flow solver and analyzed differences between these models in various idealized and physiological simulations, including exercise.

[28]: High-intensity focused ultrasound (HIFU) is used as a thermal ablation process to eliminate tumors in different body’s organs. Blood flow has a cooling effect. Conversely, ultrasounds are responsible for acoustic streaming. A three-dimensional acoustics-thermal-fluid coupling model is carried out to compute the temperature field a given hepatic cancerous region.

[29]: Imaged-based patient-specific models of the multi-branched pulmonary arteries and superior vena cava were built for five cavopulmonary connection (i.e. Glenn) patients prior to their third surgery to alleviate their congenital heart disease. Inflow and outflow boundary conditions for computational blood flow simulations (CFD) were constructed based on an iterative procedure to match available MRI and catheterization clinical data. Common trends and differences emerged from this three-dimensional CFD study; in particular low wall shear stress was found for all subjects, which is potentially deleterious. A sensitivity analysis was performed to investigate the impact of input data (clinical and modeling) to construct boundary conditions on several clinical and mechanical indicators. Among other findings, this study suggests that although 6-10% flow split imprecision seemed reasonable in terms of patient comparison, the common practice of imposing a right pulmonary artery/left pulmonary artery flow split of 55%/45% when performing patient specific simulations should be avoided.

[31]: A novel Y-shaped baffle has been proposed for the Fontan operation with promising initial results on idealized models or a single patient-specific model. The objective of this study is to comprehensively compare the hemodynamic performance and hepatic blood flow distribution of the Y-graft Fontan baffle with two current designs on multiple patient-specific models. Methods
include virtual geometrical design, computational fluid dynamics based on preoperative patient-specific data, particle tracking and sensitivity analysis, including rest and exercise conditions. The Y-graft Fontan design achieves overall superior hemodynamic performance compared with traditional designs. However, the results emphasize that designs should be customized for individual patients before clinical application.

6.3. Numerical methods for cardiac electrophysiology

Participants: Muriel Boulakia, Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Vincent Martin, Elisa Schonenone.

- [32]: We consider the problem of estimating some parameters of a model of electrocardiograms from the data of the Einthoven leads. The direct model is based on the bidomain equations in the heart and a Poisson equation in the torso. To keep the computational time reasonable, the evaluation of the direct problem is approximated with a reduced order model based on Proper Orthogonal Decomposition. The optimization problem is solved using an evolutionary algorithm. Numerical tests show that, with noisy synthetic data, the proposed procedure allows to recover ionic parameters and initial activation regions with a fair accuracy.

- [34] and [48]: In presence of a high magnetic field, the blood flow in the aorta induces an electrical potential which is responsible for an increase of the T-wave in the electrocardiogram (ECG). This phenomenon may perturb ECG-gated imaging. The aim of this numerical study is to reproduce this experimental observation through computer simulations. The proposed model consists of three components: magnetohydrodynamics (MHD) in the aorta, bidomain equations in the heart and electrical diffusion in the rest of the body. These models are strongly coupled together and solved with finite elements.

- [38]: We present an overview of our works about electrocardiogram numerical simulations.

- [45]: A reduced-order model based on Proper Orthogonal Decomposition is proposed for the bidomain equations of cardiac electrophysiology. Its accuracy is assessed through electrocardiograms in various configurations, including myocardium infarctions and long-time simulations. We show in particular that a restitution curve can efficiently be approximated by this approach. The reduced-order model is then used in an inverse problem solved by an evolutionary algorithm. Some attempts are presented to identify infarction locations from synthetic electrocardiograms.

6.4. Lung and respiration modeling

Participants: Laurent Boudin, Bérénice Grec, Muriel Boulakia, Anne-Claire Egloffé, Céline Grandmont, Ayman Moussa.

- [9]: This paper is concerned with a system that couples the incompressible Navier–Stokes equations to the Vlasov–Fokker–Planck equation. Such a system arises in the modeling of sprays, where a dense phase interacts with a disperse phase. The coupling arises from the Stokes drag force exerted by a phase on the other.

- [25]: We are concerned with the global well-posedness of a two-phase flow system arising in the modelling of fluid-particle interactions. This system consists of the Vlasov-Fokker-Planck equation for the dispersed phase (particles) coupled to the incompressible Euler equations for a dense phase (fluid) through the friction forcing.

- [49]: We obtain the Maxwell-Stefan diffusion model by studying the asymptotic behaviour of a multicomponent kinetic model when the Knudsen number goes to 0.

- [50]: We are concerned here with identifiability, stability properties and estimates for the inverse problem of identifying a Robin coefficient on some non accessible part of the boundary from available data on the other part of boundary corresponding to solutions of the Stokes equations. We first study the identifiability of Robin coefficient and then we establish a stability estimate of logarithm type using Carleman inequality.
6.5. Miscellaneous

Participant: Laurent Boudin.

- [43]: We deal with a kinetic model to describe the evolution of the opinion in a closed group where there are two opposite behaviours: conciliatory and contradictory agents. We provide an existence and uniqueness result for the model and numerically test it in some relevant cases.
SISYPHE Project-Team

5. New Results

5.1. Modeling, observation and control: systems modeled by ordinary differential equations

5.1.1. Nonlinear system identification

Participants: Pierre-Alexandre Bliman, Michel Sorine, Qinghua Zhang.

Our current researches on nonlinear system identification are mainly in the framework of the joint Franco-Chinese ANR-NSFC EBONSI project (See Section 6.5), started in March 2011 for three years, in collaboration with the Laboratory of Industrial Process Monitoring and Optimization of Peking University and with Centre de Recherche en Automatique de Nancy (CRAN). Three topics have been studied this year: system identification with a continuous time autoregressive model, system identification with quantized data, and Hammerstein-Wiener system identification.

Though discrete time models are widely used in system identification, some advantages of continuous time models are also of practical importance, in particular, the ability of fully benefiting from fast sampling devices. Our studies on this topic have resulted in a continuous time black-box model structure for nonlinear system identification, together with an efficient model estimation method. This model structure belongs to the class of continuous time nonlinear ARX (AutoRegressive with eXogenous input) models, with the particularity of being integrable. By applying techniques of adaptive observer, models of the proposed structure can be efficiently estimated from input-output data. This result has been presented at the last Journées Identification et Modélisation Expérimentale [49].

System identification is usually based on sampled and quantized data, because of the important role of digital computers. When quantized data are coded with a sufficiently large number of bits, the effect of quantization is often ignored in the design of system identification methods. However, when data are quantized with few bits, sometimes to a single bit leading to binary data, then the effect of quantization must be explicitly taken into account. Data quantization can be modeled as a non differentiable hard nonlinearity, hence the well known gradient-based optimization methods cannot be used for the identification of such nonlinear systems. We have developed a quadratic programming-based method for system identification from quantized data, which, in contrast to most existing methods, can be applied to systems with general input excitations. This result has been presented at the last IFAC World Congress [46].

A Hammerstein-Wiener system is composed of a dynamic linear subsystem preceded and followed by two static nonlinearities. Typically, the nonlinearities of such a system is caused by actuator and sensor distortions. The identification of such systems with a continuous time model had been studied by colleagues of CRAN with the refined instrumental variable (RIV) method. Stable low-pass filters were used to overcome the difficulties related to the continuous time nature of the model. Our study of this year is about the application of the Kalman filter at the place of the previously used low-pass filters. The advantages of this new method include the stability of the numerical algorithm and the fact that the Kalman filter does not color white noises.

5.1.2. Model-based fault diagnosis

Participants: Abdouramane Moussa Ali, Qinghua Zhang.

The increasing requirements for higher performance, efficiency, reliability and safety of modern engineering systems call for continuous research investigations in the field of fault detection and isolation. In the framework of the MODIPRO project funded by Paris Region ASTech, we are currently studying model-based fault diagnosis for nonlinear systems. Motivated by an application in the MODIPRO project, the considered system is modeled by nonlinear algebro-differential equations, with the particularity that the differential part of the model is linear in state variables. Instead of using general numerical solvers of algebro-differential equations, we are developing a method based on ordinary differential equation solvers, by taking advantage of the particular algebro-differential structure of the considered system.
5.2. Observation, control and traveling waves in systems modeled by partial differential equations

5.2.1. Inverse scattering for soft fault diagnosis in electric transmission lines

**Participants:** Michel Sorine, Huaibin Tang, Qinghua Zhang.

The inverse scattering theory is helpful for efficient use of the reflectometry technology in the field of electric transmission line fault diagnosis. Our recent studies on this topic have been published in [35], [33], [32]. The main progress of this year in our study has been the experimental validation of the inverse scattering-based method for soft fault diagnosis. In collaboration with Florent Loete of LGEP (Laboratoire de Génie Electrique de Paris), we have tested the inverse scattering-based method on cables used in Trucks. By slightly separating the two wires of a twisted pair following a predefined spatial profile, a soft fault in the cable is physically simulated. The spatially smoothly varying characteristic impedance of the cable is computed from the physical and geometrical parameters of the cable, and also computed from the reflection coefficient measured with a network analyzer at one end of the cable. The two results are close enough to clearly detect and to locate the physically simulated soft fault from the measured reflection coefficient. A demonstration software has been developed and registered at Agence pour la Protection des Programmes (APP). See Section 4.2.

5.2.2. Modeling of electric transmission networks

**Participants:** Mohamed Oumri, Michel Sorine, Filippo Visco Comandini, Qinghua Zhang.

The increasing number and complexity of wired electric networks in modern engineering systems is amplifying the importance of the reliability of electric connections. In the framework of the ANR 0-DEFECT project, we have studied mathematical models of complex electric networks with the aim of designing an algorithm for fault diagnosis (see [34], [20] for some theoretical results). A generalization of the Baum-Liu-Tesche (BLT) equation to the case of inhomogeneous transmission lines has been developed this year. Efforts have been made in particular to elaborate a fully automatized method for numerical simulation of complex networks with inhomogeneous transmission lines. An efficient method has been designed for the computation of the propagation matrix of each inhomogeneous transmission line and also for the computation of the scattering matrix at each network node. The implemented numerical simulator is based on these propagation and scattering matrices associated to the BLT equation, and on a numerical solution of this equation.

5.2.3. Diagnosis of insulator degradation in long electric cables

**Participants:** Leila Djaziri, Michel Sorine, Huaibin Tang, Qinghua Zhang.

For the diagnosis of insulator degradation in long electric cables, the estimation of the shunt conductance of such cables have been studied, in the framework of the ANR INSCAN project. The shunt conductance of a healthy electric cable is usually very weak. Even when the insulator in the cable is significantly degraded, the shunt conductance can still remain at a quite low level. The main difficulty in this study is due to the fact that the measurements made at the ends of a cable are hardly sensitive to the variations of the shunt conductance. To overcome this difficulty, two methods have been studied. The first one is based on the processing of long time data records. It is designed for the estimation of distributed shunt conductance, in order to detect and to locate inhomogeneous degradation of the insulator. The main idea of this method is to compensate the weak sensitivity of the measurement by long time data records. Numerical simulations have confirmed its feasibility. The second method aims at assessing the average shunt conductance along a cable. It is based on the analysis of the sensitivity of the wave propagation velocity to the shunt conductance. This method is currently tested through experiments made on cables of SNCF (Société Nationale des Chemins de Fer français).

5.3. System theory approach of some quantum systems

5.3.1. Design of strict control-Lyapunov functions for quantum systems with QND measurements

**Participants:** Hadis Amini, Mazyar Mirrahimi, Pierre Rouchon.
We have proposed a feedback scheme for quantum systems undergoing discrete-in-time non-destructive measurements. Under some observability assumptions, the proposed feedback law ensures the stabilization of any desired equilibrium state of the measurement process. This theoretical contribution has given rise to a primary conference paper [40] and we are currently working on a more complete journal paper that proves the convergence in presence of various measurement noises and uncertainties. The proposed feedback scheme has been applied in a recent experiment realized by Serge Haroche and Jean-Michel Raimond’s group at Ecole Normale Supérieure and has given rise to a journal paper [29].

5.3.2. Approximate stabilization of an infinite dimensional quantum stochastic system

Participants: Ram Somaraju, Mazyar Mirrahimi, Pierre Rouchon.

Extending our previous feedback schemes dealing only with finite dimensional quantum systems, we have proven the approximate stabilization of any desired Fock state in the microwave cavity setup of Ecole Normale Supérieure. By appropriately choosing the Lyapunov function, we avoid the mass-loss phenomena through high energy levels and ensure the pre-compactness (in probability) of the trajectories. This work has given rise to a conference paper [48] and a submitted journal paper [56].

5.3.3. On stability of continuous-time quantum filters

Participants: Hadis Amini, Mazyar Mirrahimi, Pierre Rouchon.

We have studied the stability of quantum filters for continuous-in-time measurement processes. Indeed, we have proven that the filter between the quantum state governed by a continuous time stochastic master equation driven by a Wiener process and its associated quantum-filter state is a sub-martingale. This result has given rise to a conference paper [41].

5.4. Modeling, observation and control in biosciences - Reproductive system

5.4.1. Numerical simulation of the selection process of the ovarian follicles

Participants: Benjamin Aymard, Frédérique Clément.

Collaboration with Frédéric Coquel and Marie Postel.

We have designed and implemented a numerical method to simulate a multiscale model describing the selection process in ovarian follicles [9], [8]. The PDE model consists in a quasi-linear hyperbolic system of large size, namely \( N_f \times N_f \), ruling the time evolution of the cell density functions of \( N_f \) follicles (in practice \( N_f \) is of the order of a few to twenty). These equations are weakly coupled through the sum of the first order moments of the density functions. The time-dependent equations make use of two structuring variables, age and maturity, which play the roles of space variables. The problem is naturally set over a compact domain of \( \mathbb{R}^2 \). The formulation of the time-dependent controlled transport coefficients accounts for available biological knowledge on follicular cell kinetics. We introduce a dedicated numerical scheme that is amenable to parallelization, by taking advantage of the weak coupling. Numerical illustrations assess the relevance of the proposed method both in term of accuracy and HPC achievements [50], [51].

5.4.2. Multiscale modeling of follicular ovulation as a mass and maturity dynamical system

Participants: Frédérique Clément, Philippe Michel, Danielle Monniaux.
We have analyzed the dynamics of the solutions using bifurcation tools on a reduced, ODE model [73]. In a first stage, the 2D PDE model is reduced to a 1D PDE model, where the only remaining variable is the age. This reduction is based on a result of exponential convergence in maturity; we have proved that the granulosa cell density of each follicle converges to a "dirac mass in maturity", which can be understood as: "the follicle becomes uniform in maturity". The proof is based on the crucial decay property of the maturity speed rate with respect to the maturity variable, so that the support of the cell density of each follicle concentrates its mass around a curve given by a characteristic equation. In a second stage, the mitosis rate is averaged in age, reducing the 1D PDE to a simpler system of two coupled nonlinear ODE, where each follicle is characterized by its cell number (the follicle mass) and global maturity. These variables correspond respectively to the zero-order moment and first-order moment in maturity of the cell density in the original model. The dynamics of one given follicle can then be studied with respect to the pressure exerted collectively by all other growing follicles, in the framework of dynamical games. In some sense, the pressure can be considered as an exogenous parameter, so that we can detect dynamical bifurcations according to the pressure value. Each follicle plays against the others for its survival. In the course of its terminal development, a follicle first remains in the proliferative zone of the mass-maturity plane and then enters the differentiated zone. At the transition from proliferation to differentiation, the follicle is highly sensitive to the pressure. In the worst (doomed) case, the follicle becomes atretic, due to prolonged cell loss. In the best (saved) case, it manages to go through the vulnerability zone and becomes insensitive to the pressure of other follicles.

5.4.3. Optimal control for a conservation law modeling the development of ovulation

Participants: Frédérique Clément, Peipei Shang.

Collaboration with Jean-Michel Coron

We are now investigating control problems associated to the multiscale model of follicle selection. The conditions for the triggering of the ovulatory surge, coupled with the sorting of the ovulatory follicles, define a complex, nested reachability problem. We have considered a more tractable version of that problem, which is centered on defining the optimal local control corresponding to a single ovulatory trajectory. Under some simplifying assumptions (no loss term and constant aging velocity), we have obtained analytical and numerical results in the case when the density is idealized by one or several Dirac mass. We are extending our results to the PDE original model.

5.4.4. Multiscale analysis of mixed-mode oscillations in a phantom bursting model

Participants: Frédérique Clément, Mathieu Desroches, Maciej Krupa, Alexandre Vidal.

We have carried on the study of our fast-slow model of the GnRH (gonadotropin-releasing hormone) pulse and surge generator [5], [4]. If we relax a little the constraints imposed by the biological specifications on the parameters, very rich and complex behaviours can be further exhibited by the model. More precisely, both a delay to the surge and a post-surge pause (before pulsatility resumption) may occur. A detailed examination of the pause has revealed that it is shaped by mixed-mode oscillations (MMO). We are currently investigating how the precise sequence of MMO is determined by the global return map from the surge to the pulse regime.

5.4.5. Transient synchronization of calcium oscillations in cultures of GnRH neurons.

Participants: Frédérique Clément, Maciej Krupa, Alexandre Vidal.

We have started to study the individual dynamics of GnRH neurons and the conditions under which they may synchronize. We are more specifically tackling the issue of synchronization of calcium oscillations in cultures obtained from the olfactory placodes of rhesus monkey embryos [81]. We have introduced a class of models explaining the synchronization events; their main idea was to introduce a global variable controlling the onset of synchronization that was subsequently reset by the subsequent high firing rate caused by the activation of an adaptation current.

5.5. Clinical and physiological applications

5.5.1. DynPeak: An algorithm for pulse detection and frequency analysis in hormonal time series

Participants: Frédérique Clément, Claire Médigue, Alexandre Vidal, Qinghua Zhang.
The endocrine control of the reproductive function is often studied from the analysis of luteinizing hormone (LH) pulsatile secretion by the pituitary gland. Whereas measurements in the cavernous sinus cumulate anatomical and technical difficulties, LH levels can be easily assessed from jugular blood. However, plasma levels result from a convolution process due to clearance effects when LH enters the general circulation. Simultaneous measurements comparing LH levels in the cavernous sinus and jugular blood have revealed clear differences in the pulse shape, the amplitude and the baseline. Besides, experimental sampling occurs at a relatively low frequency (typically every 10 min) with respect to LH highest frequency release (one pulse per hour) and the resulting LH measurements are noised by both experimental and assay errors. As a result, the pattern of plasma LH may be not so clearly pulsatile. Yet, reliable information on the InterPulse Intervals (IPI) is a prerequisite to study precisely the steroid feedback exerted on the pituitary level. Hence, there is a real need for robust IPI detection algorithms. We have designed an algorithm for the monitoring of LH pulse frequency, basing ourselves both on the available endocrinological knowledge on LH pulse (shape and duration with respect to the frequency regime) and synthetic LH data generated by a simple model [ 57 ]. We make use of synthetic data to make clear some basic notions underlying our algorithmic choices. We focus on explaining how the process of sampling affects drastically the original pattern of secretion, and especially the amplitude of the detectable pulses. We then describe the algorithm in details and perform it on different sets of both synthetic and experimental LH time series. We further comment on how to diagnose possible outliers from the series of IPIs which is the main output of the algorithm.
ARLES Project-Team

6. New Results

6.1. Introduction

The ARLES project-team investigates solutions in the forms of languages, methods, tools and supporting middleware to assist the development of distributed software systems, with a special emphasis on mobile distributed systems enabling the ambient intelligence/pervasive computing vision. Our research activities in 2011 have focused on the following areas:

- Dynamic interoperability among networked systems toward making them eternal, by way of on-the-fly generation of connectors based on adequate system models (§ 6.2);
- Pervasive service-oriented software engineering, focusing on supporting service composition in an increasingly heterogeneous and dynamic networking environment, while enforcing quality of service (§ 6.3);
- Service oriented middleware for the ultra large scale future Internet of Things (§ 6.4);
- Abstractions for enabling domain experts to easily compose applications on the Internet of Things (§ 6.5); and
- System-level support for application development in the context of mobile social ecosystems, while taking into account privacy, performance, and data interoperability (§ 6.6).

6.2. Emergent Middleware Supporting Interoperability in Extreme Distributed Systems

Participants: Emil Andriescu, Nelly Bencomo, Amel Bennaceur, Luca Cavallaro, Nikolaos Georgantas, Sneha-Sham Godbole, Valérie Issarny, Rachid Saadi, Daniel Sykes.

Interoperability is a fundamental challenge for today’s extreme distributed systems. Indeed, the high-level of heterogeneity in both the application layer and the underlying infrastructure, together with the conflicting assumptions that each system makes about its execution environment hinder the successful interoperation of independently developed systems. A wide range of approaches have been proposed to address the interoperability challenge [31]. Solutions that require performing changes to the systems are usually not feasible since the systems to be integrated may be legacy systems, COTS (Commercial Off-The-Shelf) components or built by third parties; neither are the approaches that prune the behavior leading to mismatches since they also restrict the systems’ functionality. Therefore, many solutions that aggregate the disparate systems in a non-intrusive way have been proposed. These solutions use intermediary software entities, called mediators, to interconnect systems despite disparities in their data and/or interaction models by performing the necessary coordination and translations while keeping them loosely-coupled. However, creating mediators requires a substantial development effort and a thorough knowledge of the application-domain, which is best understood by domain experts. Moreover, the increasing complexity of today’s distributed systems, sometimes referred to as Systems of Systems, makes it almost impossible to develop ‘correct’ mediators manually. Therefore, formal approaches are used to synthesize mediators automatically.
In light of the above, we have introduced the notion of *emergent middleware* for realizing mediators. Our research on enabling emergent mediators is done in collaboration with our partners of the CONNECT project (§ 7.1.1). Our work during the year has more specifically focused on:

- **Supporting architecture.** We have been working together with our partners in the CONNECT project on the refinement of an overall architecture supporting emergent middleware, from the discovery of networked systems to the learning of their respective behavior, and synthesis of emergent middleware enabling them to interoperate [30].

- **Affordance inference.** We have proposed an ontology-based formal model of networked systems based on their affordances, interfaces, behavior, and non-functional properties, each of which describes a different facet of the system [2]. However, legacy systems do not necessarily specify all of the aforementioned facets. Therefore, we are currently exploring techniques to infer the affordance by using textual descriptions of the interface of networked systems. More specifically, we rely on machine learning techniques to automate the inference of the affordance from the interface description by classifying the natural-language text according to a predefined ontology of affordances [17].

- **Mediator synthesis for emergent connectors.** We focus on systems that have compatible functionality, i.e., semantically matching affordances, but are unable to interact successfully due to mismatching interfaces or behaviors. We propose two approaches to enable communication between such systems:
  1. A *mapping based* approach, whose goal is to automatically synthesize a mediator model that ensures their safe interaction, i.e., deadlock-freedom and the absence of unspecified receptions. Our approach combines semantic reasoning and constraint programming to identify the semantic correspondence between networked systems’ interfaces, i.e., *interface mapping*. Unlike existing approaches that only tackle the one-to-one correspondence between actions, this approach handles the more general cases of one-to-many and many-to-many mappings.
  2. A *goal based* approach, which enables the communication of two networked systems, so that the communication satisfies a given user goal. It aligns their actions using ontology matching. The aligned processes as well as the user goal are encoded as a satisfiability problem. It relies on model checking to determine if a feasible communication trace exists that satisfies the user goal. The model checking process is reiterated so as to discover all the feasible satisfying traces, which are finally concatenated to build the mediator.

The feasibility of both of our approaches has been demonstrated through prototype tools and real-world scenarios involving heterogeneous systems.

- **Mediator synthesis for streaming connectors.** In the context of dynamic mediator synthesis, we have targeted the domain of mobile multimedia streaming, resulting in a first step that statically solves the hard problem of streaming interoperability across heterogeneous smartphone multimedia platforms. With the recent evolution of mobile phones, multimedia streaming is now commonly used in smartphones for purposes such as video broadcast, video conferencing and place shifting, which in turn highlights the importance of multimedia enabled applications. However, peer-to-peer solutions are difficult to implement because of increased node heterogeneity and their low processing power. Furthermore, existing mobile platforms such as Android, iOS, Blackberry and Windows Phone 7 support multimedia streaming (as resource consumers) either through platform specific APIs or system services. However, they use heterogeneous protocols and data formats, thus compromising interoperability.

Given the challenges above, we designed AmbiStream [11], a lightweight middleware for heterogeneous mobile devices, capable of “on the fly” adaptation. AmbiStream relies on the highly-optimized multimedia software stacks provided by smartphone platforms and adds the necessary
layers to solve interoperability. More specifically, the middleware targets: a) Streaming of prerecorded or live audio/video using an intermediary real-time protocol; b) Managing streaming protocol translation and multimedia container format adaptation to the ones supported natively by each device; and c) Extensibility in order to support new multimedia streaming protocols and multimedia container formats given its plug-in based architecture. We have used a model-driven approach to generate multi-platform plug-ins from higher level descriptions in the form of a Domain Specific Language (DSL). The defined DSL takes into account multimedia specific operations such as timing, fragmenting, multiplexing, congestion control and buffering.

- **Models@run.time.** We have recently integrated the notion of Models@run.time \(^{11}\) in our research towards emergent middleware. We use Models@run.time to extend the applicability of models and abstractions to the runtime environment. As is the case for software development models, a run-time model is often created to support reasoning. However, in contrast to development models, run-time models are used to reason about the operating environment and runtime behavior, and thus these models must capture abstractions of runtime phenomena. Different dimensions need to be balanced, including resource-efficiency (time, memory, energy), context-dependency (time, location, platform), as well as personalization (quality-of-service specifications, profiles). The hypothesis is that because Models@run.time provide meta-information for these dimensions during execution, run-time decisions can be facilitated and better automated. Thus, we anticipate that Models@run.time will play an integral role in the management of extremely distributed systems.

Our work on the use of Models@run.time has two aspects:

- We have used Models@run.time to tackle the crucial problem of uncertainty in extremely distributed systems that are aware of their own requirements. Requirements awareness helps optimize requirements satisfaction when factors that were uncertain at design time are resolved at runtime. Using our approach, we are able to maintain goal-based models in memory while the system is running. The executing system, therefore, is able to introspect and consult it goals during runtime. Crucially, at runtime we use the notion of claims to represent assumptions that cannot be verified with confidence at design time. Such claims are attached to the goal-based runtime models. By monitoring claims at runtime, their veracity can be tested. If falsified, the effect of claim negation can be propagated to the system’s goal model and an alternative means of goal realization can be selected automatically, allowing the dynamic adaptation of the system to the prevailing environmental context \(^{14},^{15},^{16}\).

- In a complementary way to the mediator synthesis approaches discussed above, we further promote the use of Models@run.time to support the runtime synthesis of software that will be part of the executing system. Specifically, we focus on the use of runtime models to support the realization of emergent middleware, i.e., the synthesis of mediators that define a sequences of actions to translate semantic actions of one system developed using a particular middleware protocol to the semantic actions of another system developed using an alternate middleware built with no prior knowledge on the former. Discovery and learning enablers capture the required knowledge of the context and environment during runtime. Supported by that knowledge, a runtime model of the mediator-to-be is reified. Reification means that the knowledge is explicitly formulated and made available for computational manipulation. The form of the runtime models is based on labeled transition systems (LTSs) which offer the behavioral semantics needed to model the interaction protocols. Ontologies complement the LTSs providing semantic reasoning about the mapping between protocols. Specifically the LTS of each protocol is annotated using ontologies to support the subsequent mapping between the protocols. From the LTS-based runtime models, mediators are synthesized.

11 Models@run.time Dagstuhl Seminar, http://www.dagstuhl.de/en/program/calendar/semhp/?semnr=11481
6.3. Revisiting the Abstractions of Service Oriented Computing for the Future Internet

Participants: Mohammad Ashiqur Rahaman, Dionysis Athanasopoulos, Sandrine Beauche, Nebil Ben Mabrouk, Nikolaos Geogantas, Valérie Issarny.

A software architecture style characterizes, via a set of abstractions, the types of: components (i.e., units of computation or data stores), connectors (i.e., interaction protocols) and possibly configurations (i.e., system structures) that serve to build a given class of systems. As such, the definition of a software architectural style is central toward eliciting appropriate design, development and runtime support for any family of systems. The service oriented architecture style may then be briefly defined as follows: (1) components map to services, which may be refined into consumer, producer or prosumer services; (2) connectors map to traditional client-service interaction protocols; and (3) configurations map to compositions of services through (service-oriented) connectors, e.g., choreography and orchestration structures. While the service-oriented architecture style is well suited to support the development of Internet-based distributed systems, it is largely challenged by the Future Internet that poses new demands in terms of sustaining qualities such as scalability, heterogeneity, mobility, awareness & adaptability that come in extreme degrees compared to the current Internet. Therefore, we have been working on eliciting software architectural abstractions for the Future Internet by building upon the service-oriented architecture style, as well as on applying them to system design, development and execution.

Complex distributed applications in the Future Internet will be to a large extent based on the open integration of extremely heterogeneous systems, such as lightweight embedded systems (e.g., sensors, actuators and networks of them), mobile systems (e.g., smartphone applications), and resource-rich IT systems (e.g., systems hosted on enterprise servers and Cloud infrastructures). These heterogeneous system domains differ significantly in terms of interaction paradigms, communication protocols, and data representation models, provided by supporting middleware platforms. Specifically considering interaction paradigms, the client/server (CS), publish/subscribe (PS), and tuple space (TS) paradigms are among the most widely employed ones today, with numerous related middleware platforms. In light of the above, we have aimed at eliciting abstractions that (i) leverage the diversity of interaction paradigms associated with today’s and future complex distributed systems, as well as (ii) enable cross-paradigm interaction to sustain interoperability in the highly heterogeneous Future Internet [19].

Existing cross-domain interoperability efforts are based on bridging communication protocols, wrapping systems behind standard technology interfaces, and/or providing common API abstractions. In particular, such techniques have been applied by the two widely established system integration paradigms, that is, service oriented architecture (SOA) and enterprise service bus (ESB). However, state of the art interoperability efforts do not or only poorly address interaction paradigm interoperability. Indeed, systems integrated via SOA and ESB solutions have their interaction semantics transformed to the CS paradigm. Then, potential loss of interaction semantics can result in suboptimal or even problematic system integration. To overcome the limitation of today’s ESB-based connectors for cross-domain interoperability in the Future Internet, we introduce a new connector type, called GA connector, which stands for “Generic Application connector”. The proposed connector type is based on the service bus paradigm in that it achieves bridging across heterogeneous connector types. However, the behavior of the GA connector type differs from that of classical ESB connectors by bridging protocols across heterogeneous paradigms, which is further realized by paying special attention to the preservation of the semantics of the composed protocols. Indeed, the GA connector type is based on the abstraction and semantic-preserving merging of the common high-level semantics of base interaction paradigms.

Eliciting Interaction Paradigm Abstractions: We introduce a systematic abstraction of interaction paradigms with the following features:

- First, we introduce base CS, PS and TS connector types, which formally characterize today’s core interaction paradigms. The proposed types comprehensively cover the essential semantics of the considered paradigms, based on a thorough survey of the related literature and representative
Then, we further abstract these connector types into a single higher-level one, the GA connector type. GA is a comprehensive connector type based on the abstract union of CS, PS, and TS, where precise identification of the commonalities or similarities between the latter has enabled the optimization of the former. Further, GA preserves by construction the semantics of CS, PS, and TS.

In more detail, connector types are formally specified in terms of: (i) their API (Application Programming Interface), and (ii) their roles, i.e., the semantics of interaction of the connected component(s) with the environment via the connector. Regarding the latter, the behavioral specification of roles from a middleware perspective relates to specifying the production and consumption of information in the network, while the semantics of the information are abstracted and dealt with at the application layer. The behaviors of the connector roles are then specified using Labeled Transition Systems (LTS). We precisely define the mapping of the roles implemented by the base connector types to/from the corresponding roles of the GA connector type.

For both the above abstraction transformations, we provide counterpart concretizations, which enable transforming GA connector primitives to CS, PS, or TS connector primitives and then to concrete middleware platforms primitives.

Furthermore, based on the GA abstraction, we introduce mapping transformations between any pair from the set \{CS, PS, TS\} via GA. The fine knowledge of CS, PS, and TS semantics, as embedded in GA, enables these mappings to be precise: differing semantics are mapped to each other in such a way that loss of semantics is limited to the minimum. These mappings relate to the definition of the glue process implemented by the GA connector, which defines how a pair of producer and consumer roles coordinates in the environment. The GA glue reconciles consumer and producer roles that may differ with respect to time and space coupling as well as scoping. Hence, GA connectors support interactions among highly heterogeneous services of the Future Internet, and especially across domains.

eXtensible Service Bus: We apply the above connector abstractions to introduce an enhanced bus paradigm, the eXtensible Service Bus (XSB). XSB features richer interaction semantics than common ESB implementations to deal effectively with the increased Future Internet heterogeneity. Moreover, from its very conception, XSB incorporates special consideration for the cross-integration of heterogeneous interaction paradigms. When mapping between such paradigms, special attention is paid to the preservation of interaction semantics. XSB has the following features:

- XSB is an abstract bus that prescribes only the high-level semantics of the common bus protocol. The XSB common bus protocol features GA semantics.
- Heterogeneous systems can be plugged into the XSB by employing binding components that adapt between the native middleware of the deployed system and the common bus protocol. This adaptation is based on the systematic abstractions and mappings discussed above
- XSB, being an abstract bus, can have different implementations. This means that it needs to be complemented with a substrate which at least supports: (1) deployment (i.e., plugging) of various systems on the bus, and (2) a common bus protocol implementing GA semantics. With respect to the latter, we envision that a GA protocol realization may either be designed and built from scratch (still supposing at least an IP-based transport substrate) or be implemented by conveying GA semantics on top of an existing higher-level protocol used as transport carrier. The latter solution can be attractive, as it facilitates GA protocol realizations in different contexts and domains.

We have carried out an early realization of XSB on the PEtALS ESB. In particular, we addressed the workflow-based orchestration of heterogeneous systems, which is a preliminary step before dealing with peer-wise system integration. This work already provides a successful feasibility study of the XSB concept. This work comprises: (i) extending the BPEL workflow language with GA API primitives; (ii) introducing transformation between the GA-extended BPEL and the standard BPEL, which consists in encapsulating GA primitives into standard BPEL primitives and enables conveying GA semantics on top of BPEL primitives.
and subsequently on top of the common bus protocol primitives; (iii) providing templates for systematic and highly facilitated building of binding components; and (iv) introducing transformations between native system interface descriptions and GA-based interface descriptions.

6.4. Service Oriented Middleware facing the Challenges of the Internet of Things

Participants: Benjamin Billet, Nikolaos Georgantas, Sara Hachem, Valérie Issarny, Roberto Speicys Cardoso, Thiago Valladares Sabino Teixeira.

Over the years, the Internet has become the most important networking infrastructure, providing an integrated entity enabling sharing, contributing, creating, using, collaborating and integrating information and knowledge by all. As a result, the Internet is changing at fast pace and is expected to evolve into the Future Internet, i.e., service-aware and self-aware federated networks that provide built-in and integrated capabilities such as: contextualization, reliability, robustness, mobility, security, service support, and self-management of communication resources and services. In our vision, The Future Internet can be defined as the union and cooperation of the Internet of Content, Internet of Services, Internet of Things, and 3D interactive Internet, supported by an expanding network infrastructure foundation. In ARLES, we chose to pay special attention to the Internet of Things (IoT). IoT is characterized by the integration of large numbers of real-world objects (or “things”) onto the Internet, with the aim of turning high-level interactions with the physical world into a matter as simple as is interacting with the virtual world today. As such, two devices that will play a key role in the IoT are sensors and actuators. In fact, such devices are already seeing widespread adoption in the highly localized systems within our cars, mobile phones, laptops, home appliances, etc. In their current incarnation, however, sensors and actuators are used for little more than low-level inferences and basic services. This is partly due to their highly specialized domains (signal processing, estimation theory, robotics, etc.), which demand application programmers to also be domain experts, and partly due to a glaring lack of interconnectivity between the different devices. Our work within that domain was focused on three related directions:

- **Challenges related to IoT:** To prepare the ground for our research on middleware for the Internet of Things, we identified the set of challenges in the IoT, namely [10]: the large scale of the Internet of Things, heterogeneity of things, unknown and dynamic network topology, unknown data-point availability, incomplete or inaccurate metadata, and conflict resolution. The scale issue arises with the millions of devices, millions of users, large amounts of data to share and services to request. The heterogeneity of the IoT is due to the fact that the network will be composed of different types of devices from different vendors with varying sensing/actuating characteristics. The unknown dynamic network topology results from the fact that devices will be mostly mobile and their availability is unknown. A related challenge is the unknown data-point availability as things, which provide the desired measurements, may leave the network or malfunction at any time. A data point is measurement of an entity of interest at a specific time. As for metadata inaccuracy, this issue is a direct result of humans, who are prone to making errors, being the main source of metadata specification. Last but not least, conflict resolution is due to the multiple stakeholders involved in the Internet of Things.

- **Middleware Requirements for the Internet of Things:** The middleware we plan on implementing should abstract things (IoT devices) as services and support dynamic service composition. To handle the IoT challenges, the middleware should also support a probabilistic discovery approach where only a subset, instead of a whole set, of devices is selected in a way that provides a good enough answer that satisfies an application’s request [10]. However, and prior to designing the middleware architecture, we extensively surveyed the literature in order to identify research challenges for service-oriented middleware design, therefore investigating service description, discovery, access and composition in the Future Internet of services [7].

- **Ontologies for the Internet of Things:** As part of our middleware architecture, we specified a set of ontologies [20] that model real-world entities as physical concepts, along with things that measure
those entities. Further, to support a smarter service composition, we also modeled mathematical formulas and physics relations as services to substitute missing thing-based services. Those services will instead compute the value of a desired measurement of an entity of interest. Finally, we also specified an ontology that describes estimation models that can be used to estimate the value of a measurement in case of a missing data point or a missing data source. Estimation models can further be used to define probabilistic discovery functions that will be executed by the middleware.

6.5. Composing Applications in the Internet of Things

**Participants:** Iraklis Leontiadis, Pankesh Patel, Animesh Pathak.

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

An important challenge to be addressed in the domain of IoT is to enable domain experts (health-care professionals, architects, city planners, etc.) to develop applications in their fields rapidly, with minimal support from skilled computer science professionals. Similar challenges have already been addressed in the closely related fields of Wireless Sensor and Actuator Networks (WSANs) and Pervasive/Ubiquitous computing. While the main challenge in the former is the extremely large scale of the systems (hundreds to thousands of largely similar nodes, sensing and acting on the environment), the primary concern in the latter has been the heterogeneity of nodes and the major role that the user’s own interaction with these nodes plays in these systems (cf. the classic “smart home” scenario where the user interacts with a smart display which works together with his refrigerator and toaster). The upcoming field of IoT includes both WSANs as well as smart appliances, in addition to the elements of the “traditional” Internet such as Web and database servers, exposing their functionalities as Web services etc. Consequently, an ideal application development abstraction of the IoT will allow (domain expert) developers to intuitively specify the rich interactions between the extremely large number of disparate devices in the future Internet of Things.

The larger goal of our research is to propose a suitable application development framework which addresses the challenges introduced above. This will most likely be achieved by a domain specific language (DSL) that exposes specific functionalities to the domain experts. The first logical step was to construct a domain model. Towards that end, we took advantage of the CRC — Classes, Responsibility, Collaboration — technique, defining the main abstract concepts, their responsibilities, and associations that represent their relationship with each other in the IoT. Specifically, we used this technique to propose a domain model [22] that addresses the following challenges:

- **Creation of common understanding.** The different terms used by different people in the IoT domain can lead to confusion, which can be alleviated by the usage of a common lexicon, as provided by a domain model. This lexicon can then be used by researchers, system programmers, as well as domain experts.

- **Modeling invariant properties.** The domain model represents the invariant properties of the domain — concepts and relationships which do not change from one application to the other. An instance of this in the IoT domain can be the notion of a sensor attached to a device. Depending on the specific applications, the type of sensors and devices can change (e.g., a light sensor attached to a smart phone), but the inherent relationship between the types of entities they represent does not.

- **Enabling modular design.** Application needs often tend to arrive in terms of behavior, which needs to be broken down and divided among the entities in the system. A good domain model aids in this process, since the capabilities of each type of entity are clearly identified. E.g., the application requirement of “the system senses the temperature of a room and keeps it steady” can be easily broken down into an application consisting of temperature sensors, computational components, and HVAC actuators, each performing its well-known role in this sense-compute-actuate loop.
As part of a related effort with a narrower focus on the domain of sensor network macroprogramming — a technique that aims to aid the wide adoption of networked sensing by providing the domain expert the ability to specify their applications at a high level of abstraction — we have explored techniques to bring Web services in the gamut of sensor network macroprogramming. Our research addresses the challenges faced by developers of systems where sensors (e.g., RFID badge sensors in an office) interact with pre-existing larger software components exposed as Web services (e.g., the office personnel access control database). As part of our work, we have proposed extensions to the data-driven ATaG macroprogramming language using which developers can easily incorporate existing Web services in their applications.

We have incorporated our continued research 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.

6.6. Addressing Middleware Challenges in Large Scale Mobile Social Networks of the Future

Participants: Sara Hachem, Valérie Issarny, Animesh Pathak, Amir Seyedi.

With the increased prevalence of advanced mobile devices (the so-called “smart” phones), interest has grown in Mobile Social Ecosystems (MSE), where users not only access traditional online Web-based social networks using their mobile devices, but are also able to use the context information provided by these devices to further enrich their interactions. In complex mobile social ecosystems of the future, the heterogeneity of software platforms on constituent nodes, combined with their intrinsic distributed nature and heterogeneity in representation of data and context, as well as user’s privacy and trust concerns, raises the need for middleware support for the development of mobile social applications. We believe that the development of mobile social applications can be greatly simplified by the presence of middleware support. To that end, we have been working on addressing the following challenges:

- **Semantic models for mobile social ecosystems.** In order to enable re-use of data between different social applications run by the same user, we have proposed an expressive and extensible model using semantic techniques to represent MSE and the interactions possible in them. This supports semantic interoperability between separately developed applications and minimizes resource-consuming operations such as data mapping and replication.

- **Efficient decentralized storage of social data.** Instead of storing the social knowledge of the whole world with a single provider — a practice performed today by common social networks such as Facebook — which can lead to privacy issues, our research endeavors to propose a middleware using which users can store their personal knowledge in a distributed manner on the devices owned by them (e.g., smart phone, home desktop, laptop). This also allows users to provide selective access to other users based on semantically defined access control policies.

- **Socially aware policies for access control.** Since social data is private and sensitive in nature, we have proposed a policy framework [21] where the user can specify both the data to be protected as well as the relevant set of peers with access to that data in a socially-aware manner (e.g., “only let my colleagues know my location during weekdays from 9 – 5”). This policy framework can be used as a guard around the user’s knowledge base, allowing access only to authorized peers. We are also working on providing end-users an easy to use editor so as to be able to specify these socially-aware policies easily.

- **Social data extraction from existing sources.** Our research includes work in enabling users to populate their social knowledge base by extracting data from their existing repositories. We have identified two types of sources of such data. The first already contain social links such as “friendship” in addition to general information, while the second do not contain social links, but may contain information which can be correlated to infer social links (e.g., call and SMS logs). We are working on a framework where adapters can be written for the former using their API to import their data; while for the latter, inference algorithms can be used to correlate data and guess/recommend social links.
Inferring trust from proximity. In mobile social network, highly sensitive private data is at risk of being shared with unwanted peers, since users may not have any knowledge about the users they socially connect with. Trust management then appears as a promising decision support for mobile users in establishing social links. However, while the literature is rich in trust models, most approaches lack appropriate trust bootstrapping, i.e., the initialization of trust values. In [24], we address this challenge by introducing proximity-based trust initialization based on the users’ behavioral data available from their mobile devices or other types of social interactions. The proposed approach is further assessed in the context of mobile social networking using users behavioral data collected by the MIT reality mining project. Results show that the inferred trust values correlate with the self-reported survey of users relationships.

We have incorporated our research in the above areas into Yarta [25], a middleware for mobile social applications. Our prototype middleware, as discussed in §5.6, currently supports application development for laptops as well as Android-powered smart phones, providing distributed storage of semantically-modeled social knowledge guarded by a rich policy framework.
4. New Results

4.1. Understanding graph representations

4.1.1. Distributed algorithms without knowledge of global parameters

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

Many fundamental local distributed algorithms are non-uniform, that is, they assume that all nodes know good estimations of one or more global parameters of the network, e.g., the maximum degree $\Delta$ or the number of nodes $n$. In [28], we introduce a rather general technique for transforming a non-uniform algorithm into a uniform one with same asymptotic complexity.

4.1.2. Asymptotic modularity

Participants: Fabien de Montgolfier, Mauricio Soto, Laurent Viennot.

Modularity has been introduced as a quality measure for graph partitioning by Newman and Girvan. It has received considerable attention in several disciplines, especially complex systems. In order to better understand this measure from a graph theoretical point of view, we study in [32], [31] the asymptotic modularity of a variety of graph classes.

4.1.3. Internet Structure

Participants: Fabien de Montgolfier, Mauricio Soto, Laurent Viennot.

In [33], [1], we study the measurement of the Internet according to two graph parameters: treewidth and hyperbolicity.

4.1.4. Multipath Spanners

Participants: Cyril Gavoille, Quentin Godfroy, Laurent Viennot.

Motivated by multipath routing, we introduce in [23], [39] a multi-connected variant of spanners.

4.1.5. $\delta$–hyperbolicity

Participants: Victor Chepoi [CNRS LIF, University of Marseille, France], Feodor Dragan [University of Ohio, USA], Bernard Estrellon [CNRS LIF, University of Marseille, France], Michel Habib [CNRS LIAFA, University of Paris Diderot, France], Yann Vaxes [University of Florence, Italy], Yang Xiang [University of Ohio, USA].

$\delta$–Hyperbolic metric spaces have been defined by M. Gromov in 1987 via a simple 4-point condition: for any four points $u, v, w, x$, the two larger of the distance sums $d(u, v) + d(w, x), d(u, w) + d(v, x), d(u, x) + d(v, w)$ differ by at most $2\delta$. They play an important role in geometric group theory, geometry of negatively curved spaces, and have recently become of interest in several domains of computer science, including algorithms and networking. In [5] paper, we study unweighted $\delta$–hyperbolic graphs. Using the Layering Partition technique, we show that every $n$–vertex $\delta$–hyperbolic graph with $\delta \geq 1/2$ has an additive $O(\delta \log n)$–spanner with at most $O(\delta n)$ edges and provide a simpler, in our opinion, and faster construction of distance approximating trees of $\delta$–hyperbolic graphs with an additive error $O(\delta \log n)$. The construction of our tree takes only linear time in the size of the input graph. As a consequence, we show that the family of $n$–vertex $\delta$–hyperbolic graphs with $\delta \geq 1/2$ admits a routing labeling scheme with $O(\delta \log^2 n)$ bit labels, $O(\delta \log n)$ additive stretch and $O(\log^2 (4\delta))$ time routing protocol, and a distance labeling scheme with $O(\log^2 n)$ bit labels, $O(\delta \log n)$ additive error and constant time distance decoder.
4.1.6. Perfect Phylogeny

4.1.6.1. Perfect Phylogeny Is \( NP \)–Hard

**Participants:** Michel Habib [CNRS LIAFA, University of Paris Diderot, France], Juraj Stacho [University of Haifa, Israel].

We answer in the affirmative [24], to the question posed by Mike Steel as a $100 challenge: “Is the following problem \( NP \)–hard? Given a ternary phylogenetic tree \( T \) and a collection \( Q \) of quartet subtrees on \( X \), is \( T \) the only tree that displays \( Q \)” As a particular consequence of this, we show that the unique chordal sandwich problem is also \( NP \)–hard.

4.1.6.2. Compatibility of Multi-states Characters

**Participants:** Michel Habib [CNRS LIAFA, University of Paris Diderot, France], Thu-Hien To [CNRS LIAFA, University of Paris Diderot, France].

Perfect phylogeny consisting of determining the compatibility of a set of characters is known to be \( NP \)–complete. We propose in [25], a conjecture on the necessary and sufficient conditions of compatibility: Given a set \( C \) of \( r \)–states full characters, there exists a function \( f(r) \) such that \( C \) is compatible if and only if every set of \( f(r) \) characters of \( C \) is compatible. According to numerous references, \( f(2) = 2, f(3) = 3 \) and \( f(r) \geq r \). Some conjectured that \( f(r) = r \) for any \( r \geq 2 \). In this paper, we present an example showing that \( f(4) \geq 5 \). Therefore it could be the case that for \( r \geq 4 \) characters, the problem behavior drastically changes. In a second part, we propose a closure operation for chordal sandwich graphs. The later problem is a common approach of perfect phylogeny.

4.1.7. Graph sandwich

**Participants:** Arnaud Durand [CNRS LIAFA, University of Paris Diderot, France], Michel Habib [CNRS LIAFA, University of Paris Diderot, France].

Graph sandwich problems were introduced by Golumbic et al. (1994) in [12] for DNA physical mapping problems and can be described as follows. Given a property \( \Pi \) of graphs and two disjoint sets of edges \( E_1, E_2 \) on a vertex set \( V \), the problem is to find a graph \( G \) on \( V \) with edge set \( E_s \) having property \( \Pi \) and such that \( E_1 \subseteq E_s \subseteq E_2 \). In [8] paper, we exhibit a quasi-linear reduction between the problem of finding an independent set of size \( k \geq 2 \) in a graph and the problem of finding a sandwich homogeneous set of the same size \( k \). Using this reduction, we prove that a number of natural (decision and counting) problems related to sandwich homogeneous sets are hard in general. We then exploit a little further the reduction and show that finding efficient algorithms to compute small sandwich homogeneous sets would imply substantial improvement for computing triangles in graphs.

4.1.8. Diameter of Real-World Undirected Graphs

**Participants:** Pierluigi Crescenzi [University of Florence, Italy], Roberto Grossi [University of Pisa, Italy], Michel Habib [CNRS LIAFA, University of Paris Diderot, France], Lorenzo Lanzi [University of Florence, Italy], Andrea Marino [University of Florence, Italy].

In [16], we propose a new algorithm for computing the diameter of undirected unweighted graphs. Even though, in the worst case, this algorithm has complexity \( O(nm) \), where \( n \) is the number of nodes and \( m \) is the number of edges of the graph, we experimentally show (on almost 200 real-world graphs) that in practice our method works in linear time. Moreover, we show how to extend our algorithm to the case of undirected weighted graphs and, even in this case, we present some preliminary very positive experimental results.

4.1.9. Parsimonious flooding in dynamic graphs

**Participants:** Hervé Baumann [CNRS LIAFA, University of Paris Diderot, France], Pierluigi Crescenzi [University of Florence, Italy], Pierre Fraigniaud [CNRS LIAFA, University of Paris Diderot, France].
An edge-Markovian process with birth-rate $p$ and death-rate $q$ generates infinite sequences of graphs $(G_0, G_1, G_2, \ldots)$ with the same node set $[n]$ such that $G_t$ is obtained from $G_{t-1}$ as follows: if $e \notin E(G_{t-1})$ then $e \in E(G_t)$ with probability $p$, and if $e \in E(G_{t-1})$ then $e \notin E(G_t)$ with probability $q$. In [2], we establish tight bounds on the complexity of flooding in edge-Markovian 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 $t' > t$. These bounds complete previous results obtained by Clementi et al. Moreover, we also show that flooding in dynamic graphs can be implemented in a parsimonious manner, so that to save bandwidth, yet preserving efficiency in term of simplicity and completion time. For a positive integer $k$, we say that the flooding protocol is $k$-active if each node forwards an information only during the $k$ time steps immediately following the step at which the node receives that information for the first time. We define the reachability threshold for the flooding protocol as the smallest integer $k$ such that, for any source $s[n]$, the $k$-active flooding protocol from $s$ completes (i.e., reaches all nodes), and we establish tight bounds for this parameter. We show that, for a large spectrum of parameters $p$ and $q$, the reachability threshold is by several orders of magnitude smaller than the flooding time. In particular, we show that it is even constant whenever the ratio $p/(p + q)$ exceeds $\log n/n$. Moreover, we also show that being active for a number of steps equal to the reachability threshold (up to a multiplicative constant) allows the flooding protocol to complete in optimal time, i.e., in asymptotically the same number of steps as when being perpetually active. These results demonstrate that flooding can be implemented in a practical and efficient manner in dynamic graphs. The main ingredient in the proofs of our results is a reduction lemma enabling to overcome the time dependencies in edge-Markovian dynamic graphs.

4.2. Distributed computational complexities

4.2.1. Local Distributed Decision
Participants: Pierre Fraigniaud [CNRS LIAFA, University of Paris Diderot, France], Amos Korman [CNRS LIAFA, University of Paris Diderot, France], David Peleg [Weizmann Institute of Science, Israel].

Inspired by sequential complexity theory, in [20] we focus on a complexity theory for distributed decision problems. We first study the intriguing question of whether randomization helps in local distributed computing, and to what extent. Our main result provides a sharp threshold for the impact of randomization on decision hereditary problems. In addition, we investigate the impact of non-determinism on local decision, and establish some structural results inspired by classical computational complexity theory. Specifically, we show that non-determinism does help, but that this help is limited, as there exist languages that cannot be decided non-deterministically. Perhaps surprisingly, it turns out that it is the combination of randomization with non-determinism that enables to decide all languages in constant time. Finally, we introduce the notion of local reduction, and establish some completeness results.

4.2.2. Asynchronous Wait-free Decision
Participants: Pierre Fraigniaud [CNRS LIAFA, University of Paris Diderot, France], Sergio Rajsbaum [Maths. Institute, University of Mexico, Mexico], Corentin Travers [Technion, Israel].

In order to capture the core of asynchronous distributed decision model, we address in [22] the wait-free model with crash failures. The set of tasks whose input is a pair $(s, t)$ and deciding whether $t \in \Delta(s)$, i.e., whether $t$ is a valid output for $s$, has been proven to be decidable in this model.

4.2.3. Mobile Distributed Decision
Participants: Pierre Fraigniaud [CNRS LIAFA, University of Paris Diderot, France], Andrzej Pelc [UQO, University of Quebec, Canada].

In [21], we partially answer the question of decidability of any language for mobile agents in a 2D environment like telecom networks or robots. It is proven that, for every agent, verifying whether (i) he/she is alone or not and (ii) he/she is able to capture the environment, is associated with the question of pertaining to an equivalence class of a map. A positive answer helps in the non-deterministic decision for any language for mobile agent.
4.2.4. Approximating the Statistics of various Properties in Randomly Weighted Graphs

**Participants:** Yuval Emek [University of Tel Aviv, Israel], Amos Korman [CNRS LIAFA, University of Paris Diderot, France], Yuval Shavitt [University of Tel Aviv, Israel].

In [19], we consider the setting of randomly weighted graphs. Under this setting, weighted graph properties typically become random variables and we are interested in computing their statistical features. Unfortunately, this turns out to be computationally hard for some weighted graph properties albeit the problem of computing the properties per se in the traditional setting of algorithmic graph theory is tractable. For example, there are well known efficient algorithms that compute the diameter of a given weighted graph, yet, computing the expected diameter of a given randomly weighted graph is \#P-hard even if the edge weights are identically distributed. In this paper, we define a family of weighted graph properties and show that for each property in this family, the problem of computing the \(k\)'th moment (and in particular, the expected value) of the corresponding random variable in a given randomly weighted graph \(G\) admits a fully polynomial time randomized approximation scheme (FPRAS) for every fixed \(k\). This family includes fundamental weighted graph properties such as the diameter of \(G\), the radius of \(G\) (with respect to any designated vertex) and the weight of a minimum spanning tree of \(G\).

4.2.5. New bounds for the controller problem

**Participants:** Yuval Emek [University of Tel Aviv, Israel], Amos Korman [CNRS LIAFA, University of Paris Diderot, France].

In [10], we establish two new lower bounds on the message complexity of the controller problem. We first prove a simple lower bound stating that any \((M, W)\)-controller must send \(\Omega(N \log M/W + 1)\) messages. Second, for the important case when \(W\) is proportional to \(M\) (this is the common case in most applications), we use a surprising reduction from the (centralized) monotonic labeling problem to show that any \((M, W)\)-controller must send \(\Omega(N \log N)\) messages. In fact, under a long lasting conjecture regarding the complexity of the monotonic labeling problem, this lower bound is improved to a tight \(\Omega(N \log^2 N)\).

4.2.6. Online computation with advice

**Participants:** Yuval Emek [University of Tel Aviv, Israel], Pierre Fraigniaud [CNRS LIAFA, University of Paris Diderot, France], Amos Korman [CNRS LIAFA, University of Paris Diderot, France], Adi Rosén [CNRS LIAFA, University of Paris Diderot, France].

In [9], we consider a model for online computation in which the online algorithm receives, together with each request, some information regarding the future, referred to as advice. We are interested in the impact of such advice on the competitive ratio, and in particular, in the relation between the size \(b\) of the advice, measured in terms of bits of information per request, and the (improved) competitive ratio. In this paper we propose the above model and illustrate its applicability by considering two of the most extensively studied online problems, namely, metrical task systems (MTS) and the \(k\)-server problem.

4.2.7. Tight Bounds For Distributed MST Verification

**Participants:** Liah Kor [Weizmann Institute of Science, Israel], Amos Korman [CNRS LIAFA, University of Paris Diderot, France], David Peleg [Weizmann Institute of Science, Israel].

In [26], we establish tight bounds for the Minimum-weight Spanning Tree (MST) verification problem in the distributed setting. Specifically, we provide an MST verification algorithm that achieves simultaneously \(\tilde{O}(|E|)\) messages and \(\tilde{O}(\sqrt{n} + D)\) time, where \(|E|\) is the number of edges in the given graph \(G\) and \(D\) is \(G\)'s diameter. On the negative side, we show that any MST verification algorithm must send \(\Omega(|E|)\) messages and incur \(\tilde{O}(\sqrt{n} + D)\) time in worst case. Our upper bound result appears to indicate that the verification of an MST may be easier than its construction.

4.2.8. Distributed verification and hardness of distributed approximation

**Participants:** Atish Das Sarma [Google research, USA], Stephan Holzer [ETH, Zurich, Switzerland], Liah Kor [Weizmann Institute of Science, Israel], Amos Korman [CNRS LIAFA, University of Paris Diderot, France], Danupon Nanongkai [Nanyang Technological University, Singapore], David Peleg [Weizmann Institute of Science, Israel], Roger Wattenhofer [ETH, Zurich, Switzerland].
In [30], we initiate a systematic study of distributed verification, and give almost tight lower bounds on the running time of distributed verification algorithms for many fundamental problems such as connectivity, spanning connected subgraph, and $s-t$ cut verification. We then show applications of these results in deriving strong unconditional time lower bounds on the hardness of distributed approximation for many classical optimization problems including minimum spanning tree, shortest paths, and minimum cut. Many of these results are the first non-trivial lower bounds for both exact and approximate distributed computation and they resolve previous open questions. Moreover, our unconditional lower bound of approximating minimum spanning tree (MST) subsumes and improves upon the previous hardness of approximation bound of Elkin [STOC 2004] as well as the lower bound for (exact) MST computation of Peleg and Rubinovich [FOCS 1999].

Our result implies that there can be no distributed approximation algorithm for MST that is significantly faster than the current exact algorithm, for any approximation factor. Our lower bound proofs show an interesting connection between communication complexity and distributed computing which turns out to be useful in establishing the time complexity of exact and approximate distributed computation of many problems.

### 4.3. Peer to Peer Networks Performance

**Participants:** Fabien Mathieu, François Baccelli.

In [3], we present and discuss possible architectures for P2P systems to manage overlays that try to cope with the underlying network.

In [40], [29], we discuss theoretical performance issues that arise from using “Live Seeding”, a technique that can be employed to leverage the capacity of a P2P/Hybrid Live Streaming Systems by utilizing the capacities of idle peers.

In [38], we propose a new paradigm for P2P networks, where the bandwidth bottleneck is not the access node anymore. This new model is versatile enough to be used in the context of classical networks with congestion control, wireless networks, or semantic networks.

### 4.4. Fault Tolerance in Distributed Networks

#### 4.4.1. Verification of population protocols

**Participants:** Hugues Fauconnier, Carole Gallet-Delporte.

In [15], we address the problem of verification by model-checking of the basic population protocol (PP) model of Angluin et al. This problem has received special attention in the last two years and new tools have been proposed to deal with it. We show that the problem can be solved by using the existing model-checking tools, e.g., Spin and Prism. In order to do so, we apply the counter abstraction to get an abstraction of the PP model which can be efficiently verified by the existing model-checking tools. Moreover, this abstraction preserves the correct stabilization property of PP models. To deal with the fairness assumed by the PP models, we provide two new recipes. The first one gives sufficient conditions under which the PP model fairness can be replaced by the weak fairness implemented in Spin. We show that this recipe can be applied to several PP models. In the second recipe, we show how to use probabilistic model-checking and, in particular, Prism to take completely in consideration the fairness of the PP models. The correctness of this recipe is based on existing theorems involving finite discrete Markov chains. An abstract of this paper has been also published in [34].

#### 4.4.2. Failure Detection

**Participants:** Hugues Fauconnier, Carole Gallet-Delporte.

What does it mean to solve a distributed task? In Paxos, Lamport proposed a definition of solvability in which every process is split into a proposer that submits commands to be executed, an acceptor that takes care of the command execution order, and a learner that receives the outcomes of executed commands. The resulting perspective of computation in which every proposed command can be executed, be its proposer correct or faulty, proved to be very useful when processes take steps on behalf of each other, i.e., in simulations.
Most interesting tasks cannot be solved asynchronously, and failure detectors were proposed to circumvent these impossibilities. Alas, when it comes to solving a task using a failure detector, we cannot leverage simulation-based techniques. A process cannot perform steps of failure detector-based computation on behalf of another process, since it cannot access the remote failure-detector module.

In [17], we propose a new definition of solving a task with a failure detector in which computation processes that propose inputs and provide outputs are treated separately from synchronization processes that coordinate using a failure detector. In the resulting framework, any failure detector is shown to be equivalent to the availability of some $k$-set agreement. As a corollary, we obtain a complete classification of tasks, including ones that evaded comprehensible characterization so far, such as renaming.

Shared objects like atomic register, test-and-set, cmp-and-swap are classical hardware primitives that help to develop fault-tolerant distributed applications. In order to compare shared objects, in [41], we consider their implementations in message passing models. With the minimal failure detector for each object, we get a new hierarchy that has only two levels. This paper summarizes recent works and results on this topic.

In [7], we first define the basic notions of local and non-local tasks for distributed systems. Intuitively, a task is local if, in a system with no failures, each process can compute its output value locally by applying some local function on its own input value (so the output value of each process depends only on the process’ own input value, not on the input values of the other processes); a task is non-local otherwise. All the interesting distributed tasks, including all those that have been investigated in the literature (e.g., consensus, set agreement, renaming, atomic commit, etc.) are non-local.

In this paper we consider non-local tasks and determine the minimum information about failures that is necessary to solve such tasks in message-passing distributed systems. As part of this work, we also introduce weak set agreement — a natural weakening of set agreement — and show that, in some precise sense, it is the weakest non-local task in message-passing systems.

### 4.4.3. Adversary disagreement and Byzantine agreement

**Participants:** Hugues Fauconnier, Carole Gallet-Delporte.

At the heart of distributed computing lies the fundamental result that the level of agreement that can be obtained in an asynchronous shared memory model where $t$ processes can crash is exactly $t + 1$. In other words, an adversary that can crash any subset of size at most $t$ can prevent the processes from agreeing on $t$ values. But what about all the other $2^n - 1 - (n + 1)$ adversaries that are not uniform in this sense and might crash certain combination of processes and not others? In [6], we present a precise way to classify all adversaries. We introduce the notion of disagreement power: the biggest integer $k$ for which the adversary can prevent processes from agreeing on $k$ values. We show how to compute the disagreement power of an adversary and derive $n$ equivalence classes of adversaries.

So far, the distributed computing community has either assumed that all the processes of a distributed system have distinct identifiers or, more rarely, that the processes are anonymous and have no identifiers. These are two extremes of the same general model: namely, $n$ processes use $\ell$ different authenticated identifiers, where $1 \leq \ell \leq n$. In [18], we ask how many identifiers are actually needed to reach agreement in a distributed system with $t$ Byzantine processes.

We show that having $3\ell + 1$ identifiers is necessary and sufficient for agreement in the synchronous case but, more surprisingly, the number of identifiers must be greater than $\frac{n + 3t}{2}$ in the partially synchronous case. This demonstrates two differences from the classical model (which has $\ell = n$): there are situations where relaxing synchrony to partial synchrony renders agreement impossible; and, in the partially synchronous case, increasing the number of correct processes can actually make it harder to reach agreement. The impossibility proofs use the fact that a Byzantine process can send multiple messages to the same recipient in a round. We show that removing this ability makes agreement easier: then, $t + 1$ identifiers are sufficient for agreement, even in the partially synchronous model.
4.4.4. Fast and compact self stabilizing verification, computation, and fault detection of an MST

**Participants:** Amos Korman [CNRS LIAFA, University of Paris Diderot, France], Shay Kutten [Technion, Israel], Toshimitsu Masuzawa [Osaka University, Japan].

In [27], we address the impact of optimizing the memory size on the time complexity, and show that this carries at most a small cost in terms of time in the context of MST. Specifically, we present a self stabilizing distributed verification algorithm whose time complexity is \(O(\log^2 n)\) in synchronous networks, or \(O(\Delta \log^2 n)\) in asynchronous networks, where \(\Delta\) denotes the largest degree of a node. More importantly, the memory size at each node remains optimal - \(O(\log n)\) bits throughout the execution. This answers an open problem posed by Awerbuch and Varghese (FOCS 1991). We also show that \(\Omega(\log n)\) time is necessary if the memory size is restricted to \(O(\log n)\) bits, even in synchronous networks. We demonstrate the usefulness of our verification scheme by using it as a module in a new self stabilizing MST construction algorithm. This algorithm has the important property that, if faults occur after the construction ended, they are detected by some nodes within \(O(\log^2 n)\) time in synchronous networks, or within \(O(\Delta \log^2 n)\) time in asynchronous networks. The rest of the nodes detect within \(O(D \log n)\) time, where \(D\) denotes the diameter. Moreover, if a constant number of faults occur, then, within the required detection time above, they are detected by some node in the \(O(\log n)\) locality of each of the faults. The memory size of the self stabilizing MST construction is \(O(\log n)\) bits per node (optimal), and the time complexity is \(O(n)\). This time complexity is significantly better than the best time complexity of previous self stabilizing MST algorithms, that was \(\Omega(n^2)\) even when using memory of \(\Omega(\log^2 n)\) bits, and even without having the above localized fault detection property. The time complexity of previous algorithms that used \(O(\log n)\) memory size was \(O(n|E|)\).

4.5. Discrete Optimization Algorithms

4.5.1. Estimating Satisfiability

**Participants:** Yacine Boufkhad, Thomas Hugel.

In [4], the problem of estimating the proportion of satisfiable instances of a given CSP (constraint satisfaction problem) can be tackled through weighting. It consists in putting onto each solution a non-negative real value based on its neighborhood in a way that the total weight is at least 1 for each satisfiable instance. We define in this paper a general weighting scheme for the estimation of satisfiability of general CSPs. First we give some sufficient conditions for a weighting system to be correct. Then we show that this scheme allows for an improvement on the upper bound on the existence of non-trivial cores in 3-SAT obtained by Maneva and Sinclair (2008) to 4.419. Another more common way of estimating satisfiability is ordering. This consists in putting a total order on the domain, which induces an orientation between neighboring solutions in a way that prevents circuits from appearing, and then counting only minimal elements. We compare ordering and weighting under various conditions.

4.5.2. Eigenvectors of three term recurrence Toeplitz matrices and Riordan group

**Participant:** Dominique Fortin.

Eigenvalues of tridiagonal (including main) Toeplitz matrices are analytically known under some regular distance to the main diagonal. Any eigenvector may be easily computed then, through a backward process; instead, in [11], we give an analytical form for each component through the reciprocation of the underlied trinomial. More generally, the connection to the Riordan group follows some bilinear iterative process.

4.5.3. Piecewise Convex Maximization problems and algorithms

**Participants:** Dominique Fortin, Ider Tseveendorj.

In [14], we provide a global search algorithm for maximizing a piecewise convex function \(F\) over a compact \(D\). We propose to iteratively refine the function \(F\) at local solution \(y\) by a virtual cutting function \(p_y(\cdot)\) and to solve
\[
\max \{ \min \{ F(x) - F(y), p_y(x) \} \mid x \in D \}
\]

Instead. We call this function either a patch, when it avoids returning back to the same local solutions, or a pseudo patch, when it possibly yields a better point. It is \textit{virtual} in the sense that the role of cutting constraints is played by additional convex pieces in the objective function. We report some computational results, that represent an improvement on previous linearization based techniques.

It is well known that maximization of any difference of convex functions could be turned into a convex maximization; in \cite{13}, we aim at a piecewise convex maximization problem instead. Despite, it may seem harder, sometimes the dimension may be reduced by 1 and the local search improved by using extreme points of the closure of the convex hull of better points. We show that it is always the case for both binary and permutation problems and give, as such instances, piecewise convex formulations for the maximum clique problem and the quadratic assignment problem.

In \cite{12}, we consider mathematical programming problems with the so-called piecewise convex objective functions. A solution method for this interesting and important class of nonconvex problems is presented. This method is based on Newton’s law of universal gravitation, multicriteria optimization and Helly’s theorem on convex bodies. Numerical experiments using well known classes of test problems on piecewise convex maximization, convex maximization as well as the maximum clique problem show the efficiency of the approach.
6. New Results

6.1. Massive mobile dense wireless networks

Participants: Cédric Adjih, Aline Carneiro Viana, Emmanuel Baccelli, Philippe Jacquet, Pascale Minet, Paul Mühlethaler, Yasser Toor.

6.1.1. Executive summary

Scaling properties of mobile ad hoc network lead to an increase of global capacity when the network density increases or when the packets can be stored for a while in mobile nodes instead of being immediately retransmitted.

Gupta and Kumar have shown in 2000 that the transport capacity per node in a multihop ad hoc network decreases in $1 / \sqrt{N \log N}$, $N$ being the number of nodes in the network. Therefore the global capacity of the network increases in $\sqrt{N} / \sqrt{\log N}$. This is a surprising result since in wired network a collection of nodes connected to a single communication resource has a transport capacity that just remains constant (i.e. the average per node capacity decreases in $1/N$).

Therefore adding space to a multihop wireless network increases the capacity: this is the space capacity paradox. When nodes randomly move, it turns to be more advantageous to store packets for a while on mobile routers instead of forwarding them immediately like hot potatoes. When the mobile router moves closer to the destination, then it can delivers packets on a much smaller number of hops. Of course the delivery delay is much longer, but the network capacity also increases by slowing non urgent packets. This is the time capacity paradox: by slowing packets, nodes mobility increases network capacity. This was hinted the first time by Grossglauser and Tse in 2002.

The great challenge is to find the good protocol and tunings that allow to adjust the delivery delay from zero to infinity in order to get a continuous increase in capacity. The challenge is two-sided: one has to keep the delivery delay between reasonable bounds and one has to consider realistic mobility models.

Existing protocols for Mobile Ad Hoc Networks (MANET) are highly efficient in routing data between mobile nodes that belong to the same connected component (cf. the protocols which have received the RFC status by the manet group of IETF). What about a disconnected network where source and destination may be located in two different connected components? In this case usual routing protocols drop packet due to host unreachable as no end-to-end route exists at that time.

A simple idea is to allow the router that has no available route to the destination to keep the packet in buffer until the conditions become more appropriate for forwarding. The forwarding conditions will change because of mobility: the router can move closer to the destination so that they belong to the same connected component and the packet can be delivered.

Indeed, the network may be continuously partitionned due to high mobility, and the traditional approach to allow a mobile node to wait for the network to be fully connected (i.e. form a unique component) or to wait to be in range of the destination may lead to unacceptable delays. Furthermore, concrete applications, such as Defence and Disaster-Relief, cannot always rely on such assumptions.

Nevertheless, even if the communicating nodes may never be within the same connected component, it is important to observe that a “communication path” may be available through time using intermediate nodes that are temporarily within reach of each other while moving, hence making such networks viable for critical applications. Depending on the nature of the environment, these networks are now commonly referred as Intermittently Connected MANET and Delay Tolerant Networks.
In between stands the problem of the fully connected network that forms a single connected component, but for which maintaining full knowledge of the topology would simply make the network collapse under its huge control traffic. In fact this is the main problem that wireless network engineering has to face, in most experiments the generation of control traffic is the main source of disruption.

6.1.2. Scientific achievements

6.1.2.1. Scaling and spatial capacity in non uniform wireless networks

We found a more precise instance of Gupta- Kumar result by using a simple but realistic network model based on slotted ALOHA with Poisson traffic. It turns out that when the traffic density increases then the average node neighborhood area shrinks so that the average encircled traffic load remains constant with an analytical expression.

In their original model Gupta and Kumar assume that the traffic density is constant, which is far from realistic. However we have derived similar generalized results when the traffic density is not uniform. In this case, the heavier is the local traffic, the smaller are the local neighborhood and the larger is the number of hops needed to cross the congested region. Therefore the shortest paths (in hop number as computed by OLSR) will have a natural tendency to avoid congested area. The path tend to follow trajectory that have analogy in non linear optic with variable indices.

6.1.2.2. Time capacity and node mobility

We have defined a protocol that takes advantage of node mobility in a general way. In short the packet stay with its host router as long as the latter does not evade too fast from its next hop (computed via a shortest path protocol such as OLSR). In the way we understand “too fast” stands the tuning parameters we discussed above. There is no need to have node geographical location and to physically measure motion vector, since everything can be done via the analysis of the dynamic of neighborhood intersections. We analytically derived performance evaluation under random walk mobility models. We plan to simulate the protocol in a real mobility scenario. This algorithm has application in Intelligent Transport System.

6.1.2.3. Overhead reduction in large networks

The first limitation of multihop wireless network is the size of the overhead per node that increases linearly with the size of the network. This is a huge improvement compared to classic internet protocols which have quadratic overhead increases. Nevertheless this till limit the network size to some thousands. We have analyzed the performance of OLSR with Fisheye feature that significantly reduce the overhead with respect to distance. In theory the overhead reduction allows to network size of several order of magnitude. Anyhow the tuning of the overhead attenuation with distance must be carefully done when the network is mobile, in order to avoid tracking failure. We showed that an overhead reduction within square root of the network size achieve this goal.

An alternative way to overhead reduction is ad hoc hierarchical routing and Distributed Hashing Table. Work has just begun in this area.

6.1.2.4. Coloring in wireless networks

Coloring is used in wireless networks to improve communication efficiency, mainly in terms of bandwidth, energy and possibly end-to-end delays. Nodes access the medium according to their color. It is the responsibility of the coloring algorithm to ensure that interfering nodes do not have the same color. First, we established complexity results about the h-hop coloring problem. Second, we focused on wireless sensor networks with grid topologies. We proposed the Vector-Based Coloring Method, denoted VCM, a new method that is able to provide an optimal periodic coloring for any radio transmission range and for any h-hop coloring, h>=1. Third, we also designed OSERENA "Optimized SchEduling RoutEr Node Activity", a distributed coloring algorithm optimized for dense wireless networks.

6.1.2.5. Complexity results about the h-hop coloring problem

In the paper we published at the WMNC 2011 conference, we define the h-hop node coloring problem, with h any positive integer, adapted to two types of applications in wireless networks. We specify both general mode for general applications and strategic mode for data gathering applications. We prove that the associated decision problem is NP-complete.
6.1.2.6. Grid coloring and the Vector-Based Coloring Method

In 2011, we also focused on wireless sensor networks with grid topologies. How does a coloring algorithm take advantage of the regularity of grid topology to provide an optimal periodic coloring, that is a coloring with the minimum number of colors? We propose the Vector-Based Coloring Method, denoted VCM, a new method that is able to provide an optimal periodic coloring for any radio transmission range and for any h-hop coloring, h≥1. This method consists in determining at which grid nodes a color can be reproduced without creating interferences between these nodes while minimizing the number of colors used. We compare the number of colors provided by VCM with the number of colors obtained by a distributed coloring algorithm with line and column priority assignments. We also provide bounds on the number of colors of optimal general colorings of the infinite grid, and show that periodic colorings (and thus VCM) are asymptotically optimal. Finally, we discuss the applicability of this method to a real wireless network.

6.1.2.7. Opportunistic routing

The model of wireless networks based on dynamic graph does not well assess the real processes in a wireless network. In particular the range of transmission can greatly vary between packets, the graph keeping only the average range. Opportunistic routing consists into taking advantage of temporary extension of the transmission range in order to gain several hops.

We have strong established theoretical performance limits in opportunistic routing. The limits are based on realistic interference scenarios in slotted Aloha. We have also investigated the impact of mobility on this theoretical limits. We have designed an opportunistic routing protocol whose performance are within a small margin of the theoretical limits. We have also conducted studies to support intelligent and adaptive forwarding, which allows a good trade-off between reliability and resource-efficiency. We then design a new protocol, called GrAnt, a new prediction-based forwarding protocol for complex and dynamic delay tolerant networks (DTNs). The proposed protocol uses the Ant Colony Optimization (ACO) metaheuristic with a greedy transition rule. This allows GrAnt to select the most promising forwarder nodes or allow for the exploitation of previously found good paths. The main motivation for using ACO is to take advantage of its population-based search and the rapid adaptation of its learning framework. Considering data from heuristic functions and pheromone concentration, the GrAnt protocol includes three modules: routing, scheduling, and buffer management.

6.1.2.8. Intermittent and delay tolerant networks

We consider the problem of routing in these networks, with the sole assumption that the speed of the node mobility is less than the speed of transmitting a packet to a neighbour. We compare this problem with sound propagation in liquid. We show that various pattern of mobility and network clustering can be described by a single parameter such as the information speed propagation.

We introduce new algorithms that route a packet toward a remote destination. The different algorithms vary depending on the buffering and the capacity capabilities of the network (i.e. if one or more copies of a packet can be sent and/or be kept). All algorithms are based on link aging rumors across connected components. The packet bounces from connected components to connected components, thanks to node mobility. We establish several analytical properties using an analogy with the sound propagation in liquid where molecules creates temporary connected components where sounds travel very fast.

Previous models assumed that the propagation of information path evolves like in a dynamic Erdos-Renyi graph leading to an epidemic flooding in O(log n) or O(1), n being the number of nodes in the network. We disprove the Erdos-Renyi model by showing via space-time considerations that the set of information path from a source to a destination is in fact much smaller than the path set in the Erdos-Renyi model. This lead to a much larger minimal delay in square root of n instead of log n. This correspond to a bounded maximal information propagation speed, whose estimate depends on the mobility model and the node density, and is root of multivariate explicit Bessel formulas.
Additionally, we have also considered the problem of data collection in global sensing and intermittently connected systems while avoiding the use of costly infrastructures (e.g., 3G). Motivated by the observation that node encounters are sufficient to build a connected relationship graph, we propose to take advantage of such inherent interactions to transform some mobile devices into delegates. We use then opportunistic delegation as a data traffic offload solution by investigating two main questions: (i) How to gain insights into social mobile networking scenarios?, (ii) How to utilize such insights to design solutions to alleviate overloaded 3G networks?. Our solution leverages usage of mobile applications requiring large data transfers by channeling the traffic to a few, socially important users in the network called VIP delegates. Mobile collectors need then only to meet delegates that, in turn, are responsible for gathering data from a subset of standard producers. We first investigate several delegation strategies based on the relative importance of nodes in their social interactions. Second, by considering a prediction strategy that estimates the likelihood of two nodes meeting each other, we investigate how the delegation strategies perform on predicted traces.

6.1.2.9. Network Coding

We study network coding for multi-hop wireless networks. We focus on the case of broadcasting where one source transmits information to all nodes in the network. Our goal is energy-efficient broadcast, that is, minimizing the total number of transmissions for broadcasting to the entire network. Note that this is a different problem for the classical problem of capacity maximization ; and assuming we are far from the network capacity limit, hence in fact, we could assume interference-free transmissions.

Our previous results, they had shown that network coding (and a simple coding strategy) was able to reach optimality for asymptotically large and dense networks, with assymptotically 100 % of the received transmissions being useful (innovative). We extended the results with the combined use of connected dominating sets and network coding: we were able to quantify (and bound) the benefits of network coding in networks where the area of the network stays fixed, and only the density increases.

We have proved that the performance of wireless random network coding are optimal in the following network model: the Erdos-Renyi random graph model and the unit disk random graph model. In particular we show in the Erdos-Renyi the network coding capacity rate outperform any Connected Dominating Set strategy by a factor of order log n. In the unit disk model we gain is larger than 60%. The result is based on the analysis of the connectivity stretch ratio of the random graphs. The connectivity stretch ratio is the ratio of the smallest degree over the connectivity number, and the connectivity stretch ratio tends to one in the two graph models.

6.1.3. Collaborations

- Professor Bernard Mans, Macquarie University, Sydney, Australia,
- TREC INRIA team,
- Professors Anelise Munaretto and Myriam Regattieri Delgado from Federal Technological University of Parana (UTFPR), Brazil,
- CNRS researcher Marcelo Dias de Amorim, LIP6/UPMC, France,
- Mathias Boc, CEA LIST, France,
- Computer Science Department, Sapienza University of Rome, Italy,
- University of St. Andrews, UK,
- Professor Leila Saidane, ENSI, Tunisia.

6.2. New generation of OLSR, new services and protocols

Participants: Cédric Adjih, Aline Carneiro Viana, Emmanuel Baccelli, Thomas Clausen, Philippe Jacquet, Pascale Minet, Ichrak Amdouni, Ridha Soua, Erwan Livolant, Paul Mühlenthaler, Yasser Toor.

6.2.1. Executive summary

The user of a mobile network very quickly experience problems with quality of service: links fade, connectivity disrupts, delays accumulate.
In a wireless network, the set of neighbors which with one node can communicate depends on transmission range, and numerous factors, and in addition the transmission range is often lower than the interference range (the range within which a node prevents correct transmissions of other nodes). Thus bandwidth reservation, a crucial step of quality of service, is an important and difficult problem.

The services and protocols that need careful adaptation are

- Connectivity continuity
- Bandwidth reservation
- Delay routing
- Connectivity control
- Autoconfiguration
- Security
- Energy efficiency
- Localization

The connectivity continuity is the most important problem. Trivial in the wired world where a link failure is a rare event, it becomes problematic in the mobile world where link failure caused by mobility are frequent and normal. The first experiments of mobile ad hoc networks with regular internet protocols miserably failed simply because either the protocol was to slow to recover link failure, or when tuned appropriately was generating such a huge overhead that the network collapsed under its own weight. A new generation of routing protocols has arised that allow a suitable control of connectivity in mobile networks. Among them the Optimized Link State Routing combines the optimization of overhead for mobile networks and the full internet legacy. It naturally provides path redundancy which accelerate link failure recovery.

The most important lesson that must be retained is that most of these optimization become NP complete, which is a significant complication compared to their counterpart in the classical wired world. The reason for the NP-completeness is two-sided: on one side the co-interferences make impossible an optimization link by link, on the other side, the large dispersion of performance measurement makes simple heuristic ineffective. As an example, routing with respect to shortest delay average does not guarantee smallest probability of high delay.

Since the bandwidth is scarce, any multimedia application such as video streaming is resource demanding. For example a TV broadcast that uses a mesh network will rapidly exhaust the bandwidth if all connections are point to point. In this case multicast protocols that allows to gather all these point to point connections in a single flow is a need.

There are two classes of multicast protocols: the tree based protocols and the network coding protocols. In the first class the protocols take advantage of the relatively small size of the recipient node set. One can show equivalent results of Gupta and Kumar scaling properties but in the multicast plan when the ratio of recipient versus network size is a fundamental parameter. When this ratio tends to one the performance naturally worsen.

When the recipient set is the whole network, one can apply the network coding scheme with random packet combination. In network coding the packets are no longer isolated: relay nodes makes linear combination of packets and transmitted mixed packets. In theory the performance of network coding is better than isolated packet multicast. In practice network coding is simpler to operate does not need topology management such as spanning trees or Connectedc Dominating Set. The reason for this is highly non intuitive, as if packet superposition was acting like state superposition in quantum mechanic, leading to non expected results.

Quality of service has become the central requirement that users expect from a network. High throughput, service continuity are critical issue for multimedia application over the wireless internet where the bandwidth is more scarce than in the wired world. A significant issue in the ad-hoc domain is that of the integrity of the network itself. Routing protocols allow, according to their specifications, any node to participate in the network - the assumption being that all nodes are behaving well and welcome. If that assumption fails - then the network may be subject to malicious nodes, and the integrity of the network fails. An important security service over mobile networks is to ensure that the integrity of the network is preserved even when attacks are launched against the integrity of the network.
6.2.2. Scientific achievements

6.2.2.1. Optimized Link State Routing (OLSR)

The routing protocol OLSR is universally known in the mobile wireless community (more than 475,000 hits on Google). It has numerous implementations and is used in many wireless networks. It is a proactive protocol with full internet legacy which is based on partial topology information exchange, that non the less provide optimal path with additive metrics (such as BGP/OSPF). It is an experimental RFC within IETF and soon will become a full standard under the name OLSRv2.

6.2.2.2. OSPF extension for wireless mesh networking

Long a near-future myth, ad hoc networks are now becoming a reality as a variety of wireless mesh networks are being deployed. Wireless mesh networks are a specific kind of ad hoc network, where terminals are essentially fixed. Even in such cases, which somewhat resembles usual networks, specific routing protocols have nevertheless to be employed, to cope with the characteristics of wireless, multi-hop communications. Such characteristics include scarce bandwidth over inherently unreliable, versatile, semi-broadcast links, and absence of a central authority in general. One of the main difficulties in this context is to cope with contradictory requirements such as, on one hand, dealing with bandwidth scarcity, which typically requires decreasing control traffic, while on the other hand, dealing with unreliable, versatile links which typically requires increasing control traffic. The two prominent routing protocols that have been developed for ad hoc networks and studied over the past decade, are the IETF standards AODV and OLSR. AODV is based on a reactive scheme (i.e. on-demand flooding to discover a path to a new destination), while OLSR is based on a proactive scheme, which is essentially an optimization of link state routing (i.e. pre-provisioning of paths to all possible destinations). OLSR is to date the most deployed such protocol, as it powers numerous wireless mesh community networks that currently flourish in various cities throughout Europe and North America. Based on this experience, the integration of ad hoc networking in the "standard" networking body is going further in several directions. One direction is the IEEE 802.11s standardization effort, which uses AODV and OLSR-derived algorithms to provide wireless mesh routing capabilities below IP. Another direction, spearheaded by the IETF, is the extension of IP routing standards such as OSPF to support ad hoc routing: in this realm we recently spun RFC 5449, as well as a series of academic publications on the subject. The idea behind extending OSPF to support ad hoc networks comes from a simple observation: OSPF is algorithmically quite similar to OLSR, as both are based on a proactive, link state approach. As on the other hand OSPF is a well-understood, widely deployed, industry-standard protocol, employing it to integrate ad hoc networks with existing infrastructure is considered by users as an easy migration path.

6.2.2.3. Multi-metric routing

Quality of service involves finding routes between two nodes in the network that satisfies a number of constraints. These constraints could be the requested bandwidth, the maximum delay, the minimum loss probability, the reliability of links, etc. This problem is NP-Complete because it combines additive metrics in the optimization problem. Hipercom proposed heuristics for finding routes that respect up to four metrics when calculating routes between source and destination. Another QoS issue is the creation of models that estimate the actual value of a metric. For example, computing the available bandwidth or the transfer delay on a link, etc. is very complex in a non-deterministic medium access such as WiFi. To resolve this problem, we developed a model for estimating the available bandwidth in a wireless network. This model is based on considering interfering links in cliques, after which we provide the maximum capacity that could be deployed in a clique. We may still enhance the model by adding a scaling factor to the clique equations in order to become more accurate when compared to real measurements.

In particular we have investigated the metric based on packet delay distribution. Since propagation delays between routers are negligible, most delays occur in queueing and medium access control processing. Contrary to previous common belief there is no need of network synchronization. The objective is to proactively determine the delay in absence of packet data traffic. The estimate of delay distribution is done via analytical method. In order to keep control on quality of service flows we use source routing forwarding options.
6.2.3. Collaborations

- Many contractual collaborations:
  - MoD (QoS, security, interconnection between the OLSR and OSPF routing domains),
  - Hitachi (Vehicular applications, OLSRv2),
- Non contractual:
  - BAE (OLSRv2),
  - Deutsche Telekom Labs/TU-Berlin, Germany,

6.3. Wireless Sensor Networks

Participants: Cédric Adjih, Aline Carneiro Viana, Emmanuel Baccelli, Thomas Clausen, Philippe Jacquet, Pascale Minet, Ichrak Amdouni, Ridha Soua, Erwan Livolant, Paul Mühlethaler, Yasser Toor.

6.3.1. Executive summary

In wireless sensor networks, we focus more particularly on:
- Spatial reuse of the bandwidth,
- Routing according to a cross-layering approach,
- Security,
- Energy efficiency,
- IPv6 support.

6.3.2. Scientific achievements

6.3.2.1. Cryptographic Protocols to Fight Sinkhole Attacks on Tree-based Routing in Wireless Sensor Networks

Wireless Sensor Networks (WSN) are penetrating more and more in our daily life. As a consequence, security has become an important matter for these networks. We introduce two new cryptographic protocols of different complexity and strength in limiting network degradation caused by sinkhole attacks on tree-based routing topologies in Wireless Sensor Networks (WSNs). The main goal of both protocols is to provide continuous operation by improving resilience against, rather than detection of, these attacks. The main benefit of providing resilience is that it allows operating (or graceful degradation) in the presence of attacks. Furthermore, while resilience mechanisms do not dismiss detection mechanisms, detection mechanisms often introduce more complexity and so, more weaknesses to the system, which might not justify their benefits. More specifically our two RESIlient and Simple Topology-based reconfiguration protocols are: RESIST-1 and RESIST-0. RESIST-1 prevents a malicious node from modifying its advertised distance to the sink by more than one hop, while RESIST-0 does not allow such lying at the cost of additional complexity.

6.3.2.2. IPv6 Protocol suite for Sensor Networks

Wireless sensor networking is a key element of the Internet of Things (IoT), a substantial part of the billions of smart objects that are soon to blend into the global IP network, from actuators to home appliances, from smart meters, to smart dust. Sensor nodes are devices used for distributed and automated monitoring of various parameters such as temperature, movement, noise or radioactivity levels etc. Sensors are scattered with minimum planning with respect to their precise physical position (including the central role of the sink, if any), and the set of peers with which a sensor can directly communicate through its wireless interface may change rapidly over time due to asynchronous sleep mode strategies, fluctuations in the radio environment, device failure or mobility. Through its wireless interface, a sensor thus connects to a communication link with undetermined connectivity properties. Sensor networks are a challenge to current IP standards, since on the one hand these protocols were designed to work on wired links and on the other hand these protocols were designed to work on machines that do not have drastic constraints in terms of CPU, power capacities, and memory, as sensor nodes do. In consequence, several key standard protocols (including TCP, UDP, DHCP,
NDP, SLAAC, and OSPF) do not function correctly in this environment. Nevertheless, IPv6-based sensor networking is a viable long term goal because it would enable generic, large scale, seamless integration of millions of sensing devices using heterogeneous radio technologies, at a low cost, and in a future-proof manner. The Internet Engineering Task Force (IETF) is currently engaged into multiple efforts addressing the limitations of existing standards concerning wireless sensor IP networking. Some of the standards under construction aim at fitting IP formats, especially IPv6 formats, to direct wireless communications using low power radio technologies such as IEEE 802.15.4, which require IP format compression. Other standards in development aim at providing multi-hop wireless sensor communication with IPv6, which requires specific routing protocols, efforts in which we actively participate, prompting numerous joint publications with both industrial and academic partners.

6.3.2.3. Coloring in wireless sensor networks
Graph coloring is used in wireless networks to optimize network resources: bandwidth and energy. We focus on grid topologies that constitute regular topologies for large or dense wireless networks. We consider various transmission ranges and identify a color pattern that can be reproduced to color the whole grid with the optimal number of colors. We obtain an optimal periodic coloring of the grid for the considered transmission range. We then evaluate the performance of a 3-hop distributed coloring algorithm, called SERENA. Through simulation results, we highlight the impact of node priority assignment on the number of colors obtained for any network and grids in particular. We then compare these optimal results on grids with those obtained by SERENA and identify directions to improve SERENA.

6.3.2.4. Coloring algorithm optimized for dense wireless networks
In 2011, we also designed OSERENA “Optimized SchEduling RoutEr Node Activity”, a distributed coloring algorithm optimized for dense wireless networks. Network density has an extremely reduced impact on the size of the messages exchanged to color the network. Furthermore, the number of colors used to color the network is not impacted by this optimization. We describe the properties of the algorithm and prove its correctness and termination. Simulation results point out the considerable gains in bandwidth.

6.3.2.5. Multichannel access in wireless sensor networks
In 2011 we started a research activity on multichannel access in wireless sensor networks. A state of the art has been published at the IFIP Wireless Days 2011 Conference.

6.3.3. Collaborations
- Many contractual collaborations:
  - Hitachi (Vehicular applications, OLSRv2),
  - OCARI2 project (industrial wireless sensor network, QoS, cross layer, energy efficiency, routing, node activity scheduling),
  - SAHARA project (wireless sensor network embedded in aircrafts),
  - STIC INRIA-Tunisian Universities: the team of Prof. Leila Saidane at ENSI (Performance improvement in a wireless sensor network),
- Non contractual:
  - BAE (OLSRv2),
  - Freie Universitaet (sensor networks, DHT),
  - Deutsche Telekom Labs/TU-Berlin, Germany,
  - University of Athens, Greece.

6.4. Vehicular and mobile applications
Participants: Cédric Adjih, Emmanuel Baccelli, Thomas Clausen, Philippe Jacquet, Pascale Minet, Paul Mühlethaler, Yasser Toor.
6.4.1. Executive summary

We have the following vision: in the future mobile internet and static internet will have their core deeply intricaded. This means that mobile ad hoc networks will be attached to the core network, form extension and even be part of it. For example in disaster area, a wireless network could replace the destroyed infrastructure and help to the emergency operations.

With this perspective items such as Autoconfiguration, Security are of crucial importance. However there is a potential conflict between a large population of fixed nodes based on ancient protocol and a smaller but more dynamic population based on new protocols. In the integration both population must cooperate in an hybrid protocol.

The difficulty is to build protocols that are as dynamic and efficient as MANET protocols but can support the legacy of the old and heavy internet protocols. The challenge is nevertheless achievable, because the dynamic part of the network needs less frequent updates from the fixed part of the network. Moreover the natural abundance of resource in the fixed part of the network allows it to support the more frequent updates from the mobile part.

OLSR has been found to be the natural best candidate for this challenge since it gathers dynamic and optimization with internet legacy.

6.4.2. Scientific achievements

6.4.2.1. Military tactical networks

During year 2011, we conducted several expertises about industrial proposals dealing with OLSR use in military tactical networks.

6.4.2.2. Protocols for vehicular networks

We have achieved numerous studies and design of protocols for vehicular networks and more specifically for V2V (Vehicle-to-Vehicle) network.

First we have studied the channel occupancy induced by the OLSR proactive routing protocol used in a linear Vehicular Ad hoc Network (VANET). Unlike previous studies, which usually use simulations to evaluate the overhead, we have proposed a simple analytical model to carry out this evaluation. Moreover, we did not evaluate the total overhead induced by the routing protocol as is usually proposed, but, for a given node, the channel occupation induced by the routing protocol.

We have studied flooding techniques for safety applications in VANETs. The typical scenario is the diffusion of an alert message after a car crash in a platoon of vehicles. The packet is diffused with the pure flooding, the multipoint relay (MPR) diffusion of OLSR and a geographic aware protocol. For OLSR we have introduced a variant (Robust-MPR) to improve the reliability. Different realistic scenarios were considered and various parameters such as vehicle density, and background traffic load were scrutinized. We have shown that the Robust-MPR and the geographic aware protocol satisfy the requirements of the safety applications while using considerably less overhead than pure flooding.

We have shown that the geographic aware protocols can be improved for the diffusion of an alert message by using opportunistic routing. We have designed OB-VAN (Opportunistic Broadcast for VANets) a new protocol that uses this idea. One of the novelty of this protocol is the use of an active signalling technique in the acknowledgement procedure to select the best relay taking advantage of the reception pattern of each message. We have studied OB-VAN in a linear VANET and have shown that it outperforms the flooding for the delay and the amount of overhead. However the delivery ratio of OB-VAN may be insufficient for safety applications. This remark has led to the design of R-OB-VAN which is a reliable variant of OB-VAN. With extensive simulations, we have shown that R-OB-VAN maintains a high delivery ratio even in the presence of packet loss due to shadowing.
We have studied the performance of the Aloha scheme in linear VANETs. This analysis assumes a SINR (Signal over Interference plus Noise Ratio) based model. In this model, we have derived the probability of a successful transmission between two vehicles at a distance of R meters. We have also computed the mean throughput according to Shannon’s law. In these two models, we have optimized the two quantities directly linked to the achievable network throughput i.e., the mean packet progress and the density of transport.

Finally, we have studied the utilization of opportunistic routing and shown that this technique is also beneficial for point to point traffic. It decreases the delay and increases the throughput compared with shortest path first routing. Moreover, we have also shown that opportunistic routing for point to point traffic eases considerably the optimization of the MAC scheme e.g. the transmission probability for Aloha and the carrier sense threshold for CSMA.

6.4.3. Collaboration

We received support from MoD for this activity.
4. New Results

4.1. Algorithms: Bandwidth Allocation in Optical Networks

Participants: Christine Fricker, Philippe Robert, James Roberts.

The development of dynamic optical switching is widely recognized as an essential requirement to meet anticipated growth in Internet traffic. Since September 2009, RAP has begun an investigation into the traffic management and performance evaluation issues that are particular to this technology. A first analysis of passive optical networks used for high speed Internet access has led to the proposal of an original dynamic bandwidth allocation algorithm and to an evaluation of its traffic capacity. Our activity on optical networking is carried out in collaboration with Orange Labs with whom we have had a research contract and are currently finalizing a new one. We have also established contacts with Alcatel-Lucent Bell Labs and had fruitful exchanges with Iraj Saniee and his team on their proposed time-domain wavelength interleaved networking architecture (TWIN).

We have also analyzed the traffic capacity of wavelength division multiplexing (WDM), passive optical networks (PONs) where user stations (optical network units) are equipped with tunable transmitters. For these systems users can use any of the multiple wavelengths to transmit their data but only within the limit determined by the number of transmitters they possess. A mean field approximation is investigated to estimate the capacity of a limited-gated multiserver polling system with a limit on the number of servers a given station can use simultaneously. The approximation provides an expression for the stability limit under very general assumptions about the traffic process and system configuration.

More generally, motivated by these next generation passive optical networks, a multi-server polling system has been studied where the number of servers that can attend to a queue simultaneously is limited. The stability condition is investigated for this model under quite general assumptions. The result is proved for unlimited service policies. The paper [1] presents a conjecture for the case of limited service policies and general service limits. A simulation study shows that the stability conditions may hold.

In 2011, we have worked on bandwidth allocation in meshed networks. A first study applies the TWIN architecture for a metropolitan area network but with an original medium access control (MAC) algorithm. This algorithm is inspired by our prior work on access networks and ensures an efficient and fair allocation of bandwidth to flows between network nodes. The paper [9] describes this network architecture and presents a performance evaluation using analytical models backed up by simulations.

The TWIN architecture is not extensible to a wide area for reasons of scalability and the excessive signalling delay between geographically distant nodes. We have therefore invented a new notion of a multipoint-to-multipoint lightpath that avoids these problems. A patent application relating to this invention has been submitted. This patent is owned by Orange following the terms of our contract with them. The second patent (that simply perfects the first invention) is jointly owned since the research was performed after the end of this contract. The submitted paper [13] describes the invention and its evaluation. A major advantage demonstrated in this paper is the energy saving achieved by the use of the proposed optical technology in place of electronic routers.

Ongoing research seeks to apply this type of networking solution to data centres, on one hand, and to geographically spread tier 1 Internet carrier networks, on the other. This work is performed in collaboration with Orange Labs and will be covered by a contract that is close to being finalized. We have also participated in the preparation of a European CELTIC project proposal that includes a work package dedicated to the development and experimentation of the network proposed in [13].

4.2. Algorithms: Content-Centric Networking

Participants: Mathieu Feuillet, Christine Fricker, Philippe Robert, James Roberts, Nada Sbihi.
RAP is participating in an ANR project named CONNECT which will contribute to the definition and evaluation of a new paradigm for the future Internet: a content-centric network (CCN) where, rather than interconnecting remote hosts like IP, the network directly manages the information objects that users publish, retrieve and exchange. CCN has been proposed by Van Jacobson and colleagues at the Palo Alto Research Center (PARC). In CCN, content is divided into packet-size chunks identified by a unique name with a particular hierarchical structure. The name and content can be cryptographically encoded and signed, providing a range of security levels. Packets in CCN carry names rather than addresses and this has a fundamental impact on the way the network works. Security concerns are addressed at the content level, relaxing requirements on hosts and the network. Users no longer need a universally known address, greatly facilitating management of mobility and intermittent connectivity. Content is supplied under receiver control, limiting scope for denial of service attacks and similar abuse. Since chunks are self-certifying, they can be freely replicated, facilitating caching and bringing significant bandwidth economies. CCN applies to both stored content and to content that is dynamically generated, as in a telephone conversation, for example. RAP is contributing to the design of CCN in two main areas:

- the design and evaluation of traffic controls recognizing that TCP is no longer applicable and queue management will require new, name-based criteria to ensure fairness and to realize service differentiation;
- the design and evaluation of replication and caching strategies that realize an optimal trade-off of expensive bandwidth for cheap memory.

The team will also contribute to the development of efficient forwarding strategies and investigate economic arguments that make CCN a viable replacement for IP.

The ANR project began in January 2011 and several task meetings have taken place. We have also held meetings with PARC establishing close cooperation with them and with some participants in the NSF project "Named Data Networking". We also participated in the CCN Community meeting in Palo Alto where we presented our work on traffic control. A paper describing the proposed flow-aware approach and results of a performance evaluation has been accepted for the conference Infocom 2012 [15].

Work on the performance of caching in CCN is ongoing. We have investigated popularity distributions for various types of content and evaluated their impact on the memory bandwidth tradeoff to be realized by CCN.

4.3. Algorithms: Channel Access algorithms in wireless networks

Participants: Mathieu Feuillet, Philippe Robert.

This is a collaboration with Thomas Bonald (Telecom ParisTech) and Alexandre Proutière (Microsoft Research). In wireless networks, to share available bandwidth between users is necessary. The bandwidth can be divided in several channels (frequency division) or the users can share the whole bandwidth by transmitting in different time slots (time division). We are studying different algorithms that allow users of a wireless network to access the channel. Those algorithms must avoid collisions and use the available bandwidth in the most efficient way. More and more wireless networks are decentralized and those algorithms must be distributed. Moreover, in order to use bandwidth in an efficient way, it is necessary to take the network topology into account. Recent studies have shown that it is possible to use the available bandwidth in a distributed and efficient way without message passing.

We studied a simplified version of the 802.11 channel access algorithm: CSMA/CA (Carrier Sense Multiple Access With Collision Avoidance). We proved that this algorithm does not use the bandwidth in the most efficient way and we proposed in 2010 a modification of this algorithm that is efficient. This result has been extended to multi-channels networks in [8] and its extended version [12].

4.4. Scaling Methods: Fluid limits in wireless networks

Participant: Philippe Robert.
This is a collaboration with Amandine Veber (CMAP, École Polytechnique). The goal is to investigate the stability properties of wireless networks when the bandwidth allocated to a node is proportional to a function of its backlog. This is, in some sense, a generalization of processor-sharing policies. We have investigated the fluid limits of simple examples of star topologies when the function used is \( \log \). We have shown that, under this scaling, some new phenomenon occurs, namely that a node may stabilize for some time at some very high level while the number of jobs of other nodes decreases at some fixed rate or remains finite, i.e. lives in the neighborhood of 0. An averaging phenomenon plays an important role for the return to equilibrium.

4.5. Algorithms: Distributed Hash Table

**Participants:** Mathieu Feuillet, Philippe Robert.

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. We modeled this problem thanks to standard queues in order to understand the behavior of a single file and the global dynamic of the system. 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 have been able to study different time scales of the system with an averaging phenomenon. An article is in preparation on this subject. A more sophisticated distributed model with mean field techniques is under investigation.

On the side of this project, we notably studied the distribution of hitting times of the classical Ehrenfest and Engset models by using martingale techniques, furthermore their asymptotic behavior has been analyzed when the size of the system increases to infinity [5].

4.6. Stochastic Modeling of Biological Networks

**Participants:** Emanuele Leoncini, Philippe Robert.

This is a collaboration with Vincent Fromion from INRA Jouy en Josas, which started on October 2010.

The goal is to propose a mathematical model of the production of proteins in prokaryotes. Proteins are biochemical compounds that play a key role in almost all the cell functions and are crucial for cell survival and for life in general. In bacteria the protein production system has to be capable to produce abut 2500 different types of proteins in different proportions (from few dozens for the replication machinery to more than 100000 for certain key metabolic enzymes). Bacteria uses more than the 85\% of their resources to the protein production, making it the most relevant process in these organisms. Moreover this production system must meet two opposing problems: on one side it must provide the minimal proteins quantities in order to ensure the smooth-running of the cell, on the other side it can not choose a “overproduction policy” for all the proteins, since this would impact the global performance of the system and of the bacterium itself.

Gene expression is intrinsically a stochastic process: gene activation/deactivation occurs by means the encounter of polymerase/repressor with the specific gene, moreover many molecules that take part in the protein production act at extremely low concentrations.

We have restated mathematically the classical model using Poisson point processes. This representation, well-known in the field of queueing networks but, as far as we know, new in the gene expression modeling, allowed us to weaken few hypothesis of the existing models, in particular the Poisson hypothesis, which is well-suited in some cases, but that, in some situations, is far from the biological reality as we consider for instance the protein assemblage.
The theoretical environment of Poisson point processes has lead us to propose a new model of gene expression which captures on one side the main mechanisms of the gene expression and on the other side it tries to consider hypothesis that are more significant from a biological viewpoint. In particular we have modeled: gene activation/deactivation, mRNA production and degradation, ribosome attachment on mRNA, protein production and degradation.

We have shown how the probability distribution of the protein production and the protein lifetime may have a significant impact on the fluctuations of the number of proteins. We have obtained analytic formulas when the duration of protein assemblage and degradation follows a general probability distribution, i.e. without the Poisson hypothesis. We have used our model to compare the variances resulting by choosing different hypotheses for the probability distribution of the protein production and degradation, in particular we have hypothesize the protein assembly and degradation to be deterministic. The model has showed how, under the previous hypothesis, the variance on the number of proteins is bigger than the classical model with the Poisson hypothesis.

4.7. Stochastic networks: large bike sharing systems

Participant: Christine Fricker.

This is a collaboration with Nicolas Gast (EPFL) started in December 2010. Bike sharing systems were launched by numerous cities as a serious alternative in urban transportation, for example Velib (20 000 bikes, 1 500 stations). One of the major issues is the availability of the resources: bikes or free slots to return the bikes. These systems have become a hot topic in Operation Research but there are few studies on these stochastic networks. To our knowledge, no theoretical study of such bike sharing systems exists taking into account the limited capacity of the stations.

We modeled this system in a symmetrical case. Mean field limit theorems give the dynamic of a large system and the stationary behavior of a single station. Analytical results are obtained and convergence proved in the standard model via Lyapunov functions. It allows to find the best ratio of bikes per station and to measure the improvement of incentive mechanisms, as choosing among two stations for example, or redistribution of bikes by trucks. It is under investigation. Further results deal with heterogeneous systems. Our goal is to propose via a theoretical study and tests simple algorithms to improve the system behavior.

4.8. Stochastic networks: heterogeneity

Participants: Christine Fricker, Hanène Mohamed.

Mean field techniques applied to non-symmetrical systems are explored. It appears as a promising way to obtain analytical results on systems with clusters.

4.9. Stochastic Networks: Jackson Networks

Participant: Danielle Tibi.

Lyapunov functions and essential spectral radius of Jackson networks, joint work with I. Ignatiouk-Robert (University of Cergy-Pontoise). A family of explicit multiplicative Lyapunov functions is constructed for any stable Jackson network. Optimizing the multiplicative factor over this family provides an upper bound for the essential spectral radius of the associated Markov process. For some particular classes of Jackson networks, this upper bound coincides with a lower bound derived from large deviations arguments, thus providing the exact value of the essential spectral radius. The main example is given by Jackson networks with routing matrix having a tree structure (in the sense that for any node i, at most one other node can route its customers to i). The result also holds for other types of routing matrices (e.g. completely symmetrical), under some conditions over the different arrival and service rates.

4.10. Scaling Methods: Interaction of TCP Flows

Participant: Philippe Robert.
This is a collaboration with Carl Graham (CMAP, École Polytechnique). Mathematical modeling of data transmission in communication networks has been the subject of intense activity for some time now. For data transmission, the Internet can be described as a very large distributed system with self-adaptive capabilities to the different congestion events that regularly occur at its numerous nodes. The coexistence of numerous connections in a network with a general number of nodes has been analyzed in a previous work through a mean-field limit of a Markovian model describing the interaction of several classes of permanent connections. In [6], this line of work has been generalized to the case when connections are not permanent but can be either active (ON) when it is transmitting data along its route, or idle (OFF). This year, the analysis of dynamic arrivals and departures has been investigated. The main technical problem is that mean-field asymptotics are not anymore usable. Instead, fluid limit schemes have to be considered in a quite delicate context, random measures.
REGAL Project-Team

6. New Results

6.1. Introduction

In 2011, we focused our research on the following areas:

- distributed algorithms for large and dynamic networks,
- Complex queries over peer-to-peer networks
- Trust and reputation management on P2P networks
- dynamic adaptation of virtual machines,
- services management in large scale environments,
- Formal and practical study of optimistic replication, incorporating application semantics.
- Decentralized commitment protocols for semantic optimistic replication.

6.2. Distributed algorithms

Participants: Luciana Arantes [correspondent], Franck Petit, Maria Potop-Butucaru [correspondent], Swan Dubois, Pierre Sens, Julien Sopena.

Our current research in the context of distributed algorithms focuses on two main axes. We are interested in providing fault-tolerant and self-* (self-organizing, self-healing and self-stabilizing) solutions for fundamental problems in distributed computing. More precisely, we target the following basic blocks: mutual exclusion, resources allocation, agreement and communication primitives.

In dynamic systems we are interested in designing building blocks for distributed applications such as: failure detectors, adequate communication primitives (publish/subscribe) and overlays. Moreover, we are interested in solving fundamental problems such as synchronization, leader election, membership and naming, and diffusion of information.

6.2.1. Failure Detectors for Dynamic Systems

Since 2009, we explore a distributed computing model of dynamic networks such as (MANET or Wireless sensor networks). The temporal variations in the network topology implies that these networks can not be viewed as a static connected graph over which paths between nodes are established beforehand. Path between two nodes is in fact built over the time. Furthermore, lack of connectivity between nodes (temporal or not) makes of dynamic networks a partitionable system, i.e., a system in which nodes that do not crash or leave the system might be not capable to communicate between themselves. In 2011 we propose a new failure detector protocol which implements an eventually strong failure detectors (⋄S) on a dynamic network with an unknown membership. Failure detector is a fundamental service, able to help in the development of fault-tolerant distributed systems. Our failure detector has the interesting feature to be time-free, so that it does not rely on timers to detect failures; moreover, it tolerates mobility of nodes and message losses [41].

6.2.2. Self-* Distributed Algorithms

The main challenges of our research activity over 2011 year were to develop self-* (self-stabilizing, self-organizing and self-healing) distributed algorithms for various type of networks. Self-stabilization is a general technique to design distributed systems that can tolerate arbitrary transient faults. Since topology changes can be considered as a transient failures, self-stabilization turns out to be a good approach to deal with dynamic networks. This is particularly relevant when the distributed (self-stabilizing) protocol does not require any global parameters, like the number of nodes or the diameter of the network. With such a self-stabilizing protocol, it is not required to change global parameters in the program when nodes join or leave the system. Therefore, self-stabilization is very desirable to achieve scalability and dynamicity.

- Snap-stabilizing Committee Coordination. The classic committee coordination problem characterizes a general type of synchronization called n-ary rendezvous as follows: 

“Professors in a certain university have organized themselves into committees. Each committee has an unchanging membership roster of one or more professors. From time to time a professor may decide to attend a committee meeting; it starts waiting and remains waiting until a meeting of a committee of which it is a member is started. All meetings terminate in finite time. The restrictions on convening a meeting are as follows: (1) meeting of a committee may be started only if all members of that committee are waiting, and (2) no two committees may convene simultaneously, if they have a common member. The problem is to ensure that (3) if all members of a committee are waiting, then a meeting involving some member of this committee is convened.”

In [31], we propose two snap-stabilizing distributed algorithms for the committee coordination problem. Snap-stabilization is a versatile technique allowing to design algorithms that efficiently tolerate transient faults. Indeed, after a finite number of such faults (e.g. memory corruptions, message losses, etc), a snap-stabilizing algorithm immediately operates correctly, without any external intervention. The first algorithm maximizes the concurrency, whereas the latter maximizes the fairness.

- **Snap-stabilizing Message Forwarding.** We focus on end-to-end request and response delivery of messages that are carried over the network. This problem is also known as the message forwarding problem. It consists in the management of network resources in order to forward messages, i.e., protocols allowing messages to move from a sender to receiver over the network. Combined with an self-stabilizing routing protocol, achieving snap-stabilization for the message forwarding problem is a very desirable property because every message sent by the sender is delivered in finite time to the receiver. In other words, no message that was actually sent after the system started is lost. In [46], we present a snap-stabilizing algorithm for the message forwarding that works on a tree topology. It uses a constant number of buffers per link, mainly $2\delta + 1$ buffers by node, where $\delta$ is the degree of the node. Therefore, it is particularly well suited for large-scale and dynamic systems, e.g, overlays used in peer-to-peer systems.

- **Self-Organizing Swarms of Robots.**

  Consider a distributed system where the computing units are weak mobile robots, i.e., devices equipped with sensors and are designed to move. By weak, we mean that the robots are anonymous, autonomous, disoriented, and oblivious, i.e., devoid of (1) any local parameter (such that an identity) allowing to differentiate any of them, (2) any central coordination mechanism or scheduler, (3) any common coordinate mechanism or common sense of direction, and (4) any way to remember any previous observation nor computation performed in any previous step. Furthermore, all the robots follow the same program (uniform or homogeneous), and there is no kind of explicit communication medium. The robots implicitly “communicate” by observing the position of the others robots. Two different environments are considered: (i) the continuous two-dimensional Euclidian space wherein robot can observe, compute and move with infinite decimal precision, and (ii) the discrete model in which the space is partitioned into a finite number of locations, conveniently represented by a graph where nodes represent locations that can be sensed, and where edges represent the possibility for a robot to move from one location to the other.

  During 2011, we mainly investigated the following problems: the gathering onto the plane, and the exploration of a finite discrete environment.

  1. **Gathering.** This problem can be stated as follows: Robots, initially located at various positions, gather at the same position in finite time and remain at this position thereafter. In [20], we investigate the self-stabilizing gathering problem in the plane, that is gathering the robots deterministically with no kind of restriction on the initial configuration. In particular, robots are allowed to share same positions in the initial configuration. Strong multiplicity detection is the ability for the robots to count the exact number of robots located at a given position. We show that assuming strong multiplicity detection, it is

---

possible to solve the self-stabilizing gathering problem with \( n \) weak robots in the semi-synchronous model if, and only if, \( n \) is odd. By contrast, we show that with an odd number of robots, the problem becomes solvable. Our proof is constructive, as we present and prove a deterministic self-stabilizing algorithm for the gathering problem.

In [19], we address the gathering in the discrete environment. We prove some asymptotical time and space complexity lower bounds to solve the problem. We propose an algorithm that is asymptotically optimal in both space and round complexities. Finally, we show that most of the assumptions we made are necessary to deterministically solve the rendezvous considering our initial scenario.

2. **Graph Exploration.**

The exploration problem is to visit a discrete and finite environment by a swarm of robots. We consider two types of explorations: The **finite exploration** and the **perpetual exploration**. The former requires that \( k \) robots, initially placed at different nodes, collectively explore a graph before stopping moving forever. By “collectively” explore we mean that every node is eventually visited by at least one robot. In [73], we propose optimal (w.r.t, the number of robots) solutions for the deterministic exploration of a grid shaped network by a team of \( k \) asynchronous oblivious robots in the asynchronous and non-atomic asynchronous model. In more details, we first show that no exploration protocol exists with less than three robots for any grid with more than three nodes, less than four robots for the \((2,2)\)-Grid, and less than five robots for the \((3,3)\)-Grid. Next, we show that the problem is solvable using only 3 robots for any \((i,j)\)-Grid, provided that \( j > 3 \). Our result is constructive as we present a deterministic algorithm that performs in the non-atomic asynchronous model. We also present specific deterministic protocols for the \((3,3)\)-Grid using five robots.

The second type of exploration is the perpetual exploration. It requires every possible location to be visited by each robot infinitely often. In [32], we investigate the exclusive perpetual exploration of grid shaped networks. We focus on the minimal number of robots that are necessary and sufficient to solve the problem in general grids. In more details, we prove that three deterministic robots are necessary and sufficient, provided that the size of the grid is \( nxm \) with \( 3 \leq n \leq m \) or \( n = 2 \) and \( m \geq 4 \). Perhaps surprisingly, and unlike results for the exploration with stop problem (where grids are "easier" to explore and stop than rings with respect to the number of robots), exclusive perpetual exploration requires as many robots in the ring as in the grid. Furthermore, we propose a classification of configurations such that the space of configurations to be checked is drastically reduced. This pre-processing lays the bases for the automated verification of our algorithm for general grids as it permits to avoid combinatorial explosion.

6.2.3. **Combining Fault-Tolerance and Self-stabilization in Dynamic Systems**

Recently, we started to investigate complex faults scenarios. Distributed fault-tolerance can mask the effect of a limited number of permanent faults, while self-stabilization provides forward recovery after an arbitrary number of transient faults hit the system. FTSS (Fault-Tolerant Self-Stabilizing) protocols combine the best of both worlds since they tolerate simultaneously transient and (permanent) crash faults. To date, deterministic FTSS solutions either consider static (i.e., fixed point) tasks, or assume synchronous scheduling of the system components. We proposed in [30] a fault-tolerant and stabilizing simulation of an atomic register. The simulation works in asynchronous message-passing systems, and allows a minority of processes to crash. The simulation stabilizes in a pragmatic manner, by reaching a long execution in which it runs correctly. A key element in the simulation is a new combinatorial construction of a bounded labeling scheme accommodating arbitrary labels, including those not generated by the scheme itself. Our simulation uses a self-stabilizing implementation of a data-link over non-FIFO channels [61]. In [23] we present the first study of deterministic FTSS solutions for dynamic tasks in asynchronous systems, considering the unison problem as a benchmark.
Unison can be seen as a local clock synchronization problem as neighbors must maintain digital clocks at most one time unit away from each other, and increment their own clock value infinitely often. We present several impossibility results for this difficult problem and propose a FTSS solution (when the problem is solvable) for the state model that exhibits optimal fault containment.

6.3. Peer-to-peer systems

Participants: Pierre Sens [correspondent], Nicolas Hidalgo, Sergey Legtchenko, Olivier Marin, Sébastien Monnet, Gilles Muller, Maria Potop-Butucaru, Mathieu Valero.

6.3.1. Peer-to-peer storage

Distributed Hash Table (DHTs) provide a means to build a completely decentralized, large-scale persistent storage service from the individual storage capacities contributed by each node of the peer-to-peer overlay. However, persistence can only be achieved if nodes are highly available, that is, if they stay most of the time connected to the overlay. Churn (i.e., nodes connecting and disconnecting from the overlay) in peer-to-peer networks is mainly due to the fact that users have total control on their computers, and thus may not see any benefit in keeping their peer-to-peer client running all the time.

When connection/disconnection frequency is too high in the system, data-blocks may be lost. This is true for most current DHT-based system’s implementations. To avoid this problem, it is necessary to build really efficient replication and maintenance mechanisms. Since 2008 we study the effect of churn on an existing DHT-based P2P system namely PAST/Pastry. We have proposed RelaxDHT [25], a churn-resilient peer-to-peer DHT. RelaxDHT proposes an enhanced replication strategy with relaxed placement constraints, avoiding useless data transfers and improving transfer parallelization. This new replication strategy is able to cut down by 2 the number of data-block losses compared to PAST DHT. We are now starting to study the use of erasure coding mechanisms along with replication within DHTs. Our goal is to propose hybrid mechanisms to find a good tradeoff among 1) churn-resilience, 2) maintenance cost, and 3) storage space.

6.3.2. Overlays

Large-scale distributed systems gather thousands of peers spread all over the world. Such systems need to offer good routing performances regardless of their size and despite high churn rates. To achieve that requirement, the system must add appropriate shortcuts to its logical graph (overlay). However, to choose efficient shortcuts, peers need to obtain information about the overlay topology. In case of heterogeneous peer distributions, retrieving such information is not straightforward. Moreover, due to churn, the topology rapidly evolves, making gathered information obsolete. State-of-the-art systems either avoid the problem by enforcing peers to adopt a uniform distribution or only partially fulfill these requirements. To cope with this problem, we propose DONUT [47], a mechanism to build a local map that approximates the peer distribution, allowing the peer to accurately estimate graph distance to other peers with a local algorithm. The evaluation performed with real latency and churn traces shows that our map increases the routing process efficiency by at least 20% compared to the state-of-the-art techniques. It points out that each map is lightweight and can be efficiently propagated through the network by consuming less than 10 bps on each peer.

6.3.3. Distributed trees

Publish/Subscribe implemented on top of distributed R-trees (DR-trees) overlays offer efficient DHT-free communication primitives. We have then extend the distributed R-trees (DR-trees) in order to reduce event delivery latency in order to meet the requirements of massively distributed video games such that pertinent information is quickly distributed to all the interested parties without degrading the load of nodes neither increasing the number of noisy events. In 2011, we explore how to improve robustness of distributed trees. Since each single crash can potentially break the tree structure connectivity, DR-trees are crash-sensitive. We this have proposed a fault tolerant approach which exploits replication of non leaf nodes in order to ensure the tree connectivity in presence of crashes. This work will be published in [56].
In [85], we consider a complete binary tree and construct a random pairing between leaf nodes and internal nodes. We prove that the graph obtained by contracting all pairs (leaf-internal nodes) achieves a constant node expansion with high probability. In the context of P2P overlays our result can be interpreted as follows: if each physical node participating to the tree overlay manages a random pair that couples one virtual internal node and one virtual leaf node then the physical-node layer exhibits a constant expansion with high probability which improves the robustness of the overlay.

6.3.4. Complex query over peer-to-peer networks

A major limitation of DHTs is that they only support exact-match queries. In order to offer range queries over a DHT it is necessary to build additional indexing structures. Prefix-based indexes, such as Prefix Hash Tree (PHT), are interesting approaches for building distributed indexes on top of DHTs. Nevertheless, the lookup operation of these indexes usually generates a high amount of unnecessary traffic overhead which degrades system performance by increasing response time. In [42], we propose a novel distributed cache system called Tabu Prefix Table Cache (TPT-C), aiming at improving the performance of the Prefix-trees. We have implemented our solution over PHT, and the results confirm that our searching approach reduces up to a 70% the search latency and traffic overhead. In [44], we propose DRing an efficient layered solution that directly supports range queries over a ring-like DHT structure. We improve load balancing by using only the nodes that store data, and by updating neighbour information through an optimistic approach.

6.3.5. Trust management in peer-to-peer networks

An ongoing research work focuses on trust assessment in dynamic systems. Even if it is near impossible to fully trust a node in a P2P system, managing a set of the most trusted nodes in the system can help to implement more trusted and reliable services. Using these nodes can reduce the probability of introducing malicious nodes in distributed computations. Our work aims at the following objectives: 1. To design a distributed membership algorithm for structured Peer to Peer networks in order to build a group of trusted nodes. 2. To design a maintenance algorithm to periodically clean the trusted group so as to avoid nodes whose reputation has decreased under the minimum value. 3. To provide a way for a given node X to find at least one trusted node. 4. To design a prototype of an information system, such as a news dissemination system, that relies on the trusted group. In 2011, we propose the CORPS system for building a community of reputable peers in Distributed Hash Tables [26].

6.4. Virtual machine (VM)

Participants: Harris Bakiras, Bertil Folliot, Gaël Thomas [correspondent], Gilles Muller [correspondent], Julia Lawall, Arie Middlekoop, Thomas Preud’homme, Suman Saha.

Our research interests are in improving the way systems software is developed. One theme of our research is the development of virtual machines with a specific focus on resource management, isolation and concurrency management. Another theme of our research is related to bug finding in systems software.

6.4.1. Virtual machines

Isolation in OSGi: The OSGi framework is a Java-based, centralized, component oriented platform. It is being widely adopted as an execution environment for the development of extensible applications. However, current Java Virtual Machines are unable to isolate components from each other's. By modifying shared variables or allocating too much memory, a malicious component can freeze the complete platform. We work on I-JVM, a Java Virtual Machines that provides a lightweight approach to isolation while preserving the compatibility with legacy OSGi applications. Our evaluation of I-JVM shows that it solves the 15 known OSGi vulnerabilities due to the Java Virtual Machine with an overhead below 20%. I-JVM has been presented in DSN 2009.
VMKit: Managed Runtime Environments (MREs), such as the JVM and the CLI, form an attractive environment for program execution, by providing portability and safety, via the use of a bytecode language and automatic memory management, as well as good performance, via just-in-time (JIT) compilation. Nevertheless, developing such a fully featured MRE, including features such as a garbage collector and JIT compiler, is a herculean task. As a result, new languages cannot easily take advantage of the benefits of MREs, and it is difficult to experiment with extensions of existing MRE based languages. VMKit is a first attempt to build a common substrate that eases the development of high-level MREs. We have successfully used VMKit to build two MREs: a Java Virtual Machine (J3) and a Common Language Runtime (N3). VMKit has performance comparable to the well established open source MREs Cacao, Apache Harmony and Mono. VMKit is freely distributed under the LLVM licence with the LLVM framework developed by the University of Illinois at Urbana-Champaign and now maintained by Apple.

A third MRE is being build in cooperation with the “Algorithms, Programmes and Resolution” team in LIP6. This integrates a functional machine (the Zinc Abstract Machine) in VMKit and show that the adaptation at the language level of our virtual machine. This project has been funded by the LIP6 in 2009-10 and 2010-11.

6.4.2. Semantic patches

Open source infrastructure software, such as the Linux operating system, Web browsers and n-tier servers, has become a well-recognized solution for implementing critical functions of modern life. Furthermore, companies and local governments are finding that the use of open source software reduces costs and allows them to pool their resources to build and maintain infrastructure software in critical niche areas. Nevertheless, the increasing reliance on open source infrastructure software introduces new demands in terms of security and safety. In principle, infrastructure software contains security features that protect against data loss, data corruption, and inadvertent transmission of data to third parties. In practice, however, these security features are compromised by a simple fact: software contains bugs.

We are developing a comprehensive solution to the problem of finding bugs in API usage in open source infrastructure software based on our experience in using the Coccinelle code matching and transformation tool, and our interactions with the Linux community.

Coccinelle has been successfully used for finding and fixing bugs in systems code. One of our main recent results is an extensive study of bugs in Linux 2.6 [51] that has permitted us to demonstrate that the quality of code has been improving over the last six years, even though the code size has more than doubled.

We have used Coccinelle to generate traditional patches for improving the safety of Linux. Some Linux developers have also begun to use the tool. Over 800 patches developed using Coccinelle have been integrated into the mainline Linux kernel. As part of the ABL ANR project, we are building on the results of Coccinelle by designing semantic patches to identify API protocols and detect violations in their usage [24].

Another work done as part of the ANR ABL project, and as the topic of Suman Saha’s PhD thesis, is the improvement of error handling code in Linux. We developed a program analysis for identifying the code structures used to represent error handling code and a transformation to convert existing error handling code to use gotos to shared cleanup code, which is the style preferred by the Linux community [53]. We subsequently worked on finding bugs in error handling code, following an approach that focuses on local patterns, i.e., within the current function, rather than patterns occurring across the entire code base. This approach has a low rate of false positives and can find bugs in the use of rarely called functions [39].

6.5. Hosted database replication service

Participant: Mesaac Makpangou [correspondent].

Today, the vast majority of content distributed on the web are produced by web 2.0 applications. Examples of such applications include social networks, virtual universities, multi-players games, e-commerce web sites, and search engines. These applications rely on databases to serve end-users’ requests. Hence, the success of these applications/services depends mainly on the scalability and the performance of the database backend.
The objective of our research is to provide a hosted database replication service. With respect to end-users applications, this service offers an interface to create, to register, and to access databases. Internally, each hosted database is fragmented and its fragments are replicated towards a peer-to-peer network. We anticipate that such a service may improve the performance and the availability of popular web applications, thanks to partial replications of backend databases. Partial database replication on top of a peer-to-peer network raises a number of difficult issues: (i) enforcing replica consistency in presence of update transactions, without jeopardizing the scalability and the performance of the system? (ii) accommodating the dynamic and the heterogeneity of a peer-to-peer network with the database requirements?

We designed a database access protocol, capable to spread out a transaction’s accesses over multiple database fragments replicas while guaranteeing that each transaction observes a consistent distributed snapshot of a partially replicated database. We have also proposed a replica control substrate that permits to enforce 1-Copy SI for database fragments replicated over a wide area network. For that, unlike most database replication, we separate the synchronisation from the certification concerns.

A small-scale group of schedulers that do not hold database replicas, cooperate with one another to certify update transactions. Only certified transactions are notified to replicas. Furthermore, each replica will be notified only the transactions that impact the that it stores. Thanks to this separation, we avoid waste of computation resource at replicas that will be used to decide whether to abord or commit an update transaction; Our design choices also permit to reduce bandwidth consumption.

In 2010, we focus on the development of a prototype implementation of the complete system. The current prototype includes: a tool that helps fragment a database into fragments; the support to deploy dynamically, for each fragment, the suitable number of replicas towards the network of hosting peers; the implementation of our proposal (i.e. our database access protocol and our replica control substrate); and the JDBC API extension for accessing replicated databases.

6.6. Commitment protocols for WAN replication

Participants: Marc Shapiro [correspondent], Pierre Sutra, Masoud Saeida Ardekani.

In a large-scale distributed system, replication is an essential technique for improving availability and read performance. However, writes raise the issue of consistency, especially in the presence of concurrent updates, network failures, and hardware or software crashes. So-called consensus constitutes a major primitive to solving these issues. The performance of large-scale systems depends crucially on the latency of consensus, especially in wide-area networks; to decrease it, we focus on generalised consensus algorithms, i.e., ones that leverage the commutativity of operations and/or the spontaneous ordering of messages by the network. One such algorithms is Generalized Paxos, which does not order concurrent commutative operations. However, when a collision occurs (i.e., two replicas receive non-commuting operations in a different order) Generalized Paxos requires a very high latency to recover, completely negating the gain. We designed FGGC, a new generalised consensus algorithm that minimises the cost of recovering from a collision, without decreasing resilience to faults. FGGC achieves the optimal latency (two communication steps when processes receive non-commutative operations in the same order, and three otherwise) when there are no faults. FGGC remains optimally fault-tolerant, as it tolerates $f < n/2$ crash faults and requires only $f + 1$ processes to make progress. Our experimental evaluation of FGGC shows that it is more efficient than the competing protocols. Another topic of relevance in WANs is partial replication, i.e., where any given server holds only a fraction of all shared objects. This decreases the workload per server and improves access times. However, this makes transactional concurrency control more difficult; indeed most existing algorithms assume full replication. We designed and implemented two genuine consensus protocols for partial replication, i.e., ones in which only relevant replicas need participate in the commit of a transaction. They were evaluated experimentally above the BerkeleyDB database engine. This work is the topic of Pierre Sutra’s PhD thesis.

6.7. Optimistic approaches in collaborative editing

Participants: Marc Shapiro [correspondent], Marek Zawirski, Pierpaolo Cincilla.
In recent years, the Web has seen an explosive growth of massive collaboration tools, such as wiki and weblog systems. By the billions, users may share knowledge and collectively advance innovation, in various fields of science and art. Existing tools, such as the MediaWiki system for wikis, are popular in part because they do not require any specific skills. However, they are based on a centralised architecture and hence do not scale well. Moreover, they provide limited functionality for collaborative authoring of shared documents.

A natural research direction is to use P2P techniques to distribute collaborative documents. This raises the issue of supporting collaborative edits, and of maintaining consistency, over a massive population of users, shared documents, and sites.

In order to avoid complex and unnatural concurrency control and synchronisation, and to enable different styles of collaboration (from online “what you see is what I see” to fully asynchronous disconnected work) we invented the concept of a Commutative Replicated Data Type (CRDT). A CRDT is one where all concurrent operations commute. The replicas of a CRDT converge automatically, without complex concurrency control.

In the context of collaborative editing, we propose, a novel CRDT design called Treedoc. An essential property is that the identifiers of Treedoc atoms are selected from a dense space. We study practical alternatives for implementing the identifier space based on an extended binary tree. We also focus storage alternatives for data and meta-data, and mechanisms for compacting the tree. In the best case, Treedoc incurs no overhead with respect to a linear text buffer. We validate the results with traces from existing edit histories.

Treedoc will be used in ANR projects PROSE (Section 7.1.5 ) and STREAMS, and will be further studied and developed in ANR project ConcoRDanT (Section 7.1.3 ) and under a Google European Doctoral Fellowship.

6.8. CRDTs, a principled approach to eventual consistency

Participants: Marc Shapiro [correspondent], Marek Zawirski.

Most well-studied approaches to replica consistency maintain a global total order of operations. This serialisation constitutes a performance and scalability bottleneck, while the CAP theorem imposes a trade-off between consistency and partition-tolerance. An alternative approach, eventual consistency or optimistic replication, is attractive. A replica may execute an operation without synchronising a priori with other replicas. The operation is sent asynchronously to other replicas; every replica eventually applies all updates, but possibly in different orders. This approach ensures that data remains available despite network partitions, and is perceived to scale well and to provide acceptable quality of service. The consensus bottleneck remains but is off the critical path. However, reconciliation is generally complex. There is little theoretical guidance on how to design a correct optimistic system, and ad-hoc approaches have proven brittle and error-prone. We propose a simple, theoretically sound approach to eventual consistency, the concept of a convergent or commutative replicated data type (CRDT), for which some simple mathematical properties ensure eventual consistency. Provably, any CRDT converges to a common state that is equivalent to some sequential execution. A CRDT requires no synchronisation, thus every update can execute immediately, unaffected by network latency, faults, or disconnection. It is extremely scalable and is fault-tolerant, and does not require much mechanism. Previously, only a handful of CRDTs were known. Our current research aims to push the CRDT envelope, to study the principles of CRDTs, and to design a library of useful CRDTs. So far we have designed variations on registers, counters, sets, maps (key-value stores), graphs, and sequences. Potential application areas include computation in delay-tolerant networks, latency tolerance in wide-area networks, disconnected operation, churn-tolerant peer-to-peer computing, and partition-tolerant cloud computing. CRDTs are the main topic of ANR project ConcoRDanT (Section 7.1.3 ). This research is also funded in part by a Google European Doctoral Fellowship.
6. New Results

6.1. Design and Performance Analysis of Wireless Networks

Participants: François Baccelli, Florence Bénézit, Bartłomiej Blaszczyszyn, Chung Shue Chen, Mir Omid Haji Mirsadeghi, Frédéric Morlot, Tien Viet Nguyen, Van Minh Nguyen.

This axis bears on the analysis and the design of wireless access communication networks. Our contributions are organized in terms of network classes: cellular networks, wireless LANs and MANETs, VANETs. We also have a section on generic results that regard more general wireless networks. We are interested both in macroscopic models, which are particularly important for economic planning and in models allowing the definition and the optimization of protocols. Our approach combines several tools, queueing theory, point processes, stochastic geometry, random graphs, distributed control algorithms, self organization protocols.

6.1.1. Cellular Networks

The activity on cellular networks has several complementary facets ranging from performance evaluation to protocol design. The work is mainly based on strong collaborations with Alcatel-Lucent and Orange Labs.

6.1.1.1. Effect of Opportunistic Scheduling on the Quality of Service Perceived by the Users in OFDMA Cellular Networks

Our objective in [20] is to analyze the impact of fading and opportunistic scheduling on the quality of service perceived by the users in an Orthogonal Frequency Division Multiple Access (OFDMA) cellular network. To this end, assuming Markovian arrivals and departures of customers that transmit some given data volumes, as well as some temporal channel variability (fading), we study the mean throughput that the network offers to users in the long run of the system. Explicit formulas are obtained in the case of allocation policies, which may or may-not take advantage of the fading, called respectively opportunistic and non-opportunistic. The main practical results of the present work are the following. Firstly we evaluate for the non-opportunist allocation the degradation due to fading compared to Additive White Gaussian Noise (AWGN) (that is, a decrease of at least 13% of the throughput). Secondly, we evaluate the gain induced by the opportunistic allocation. In particular, when the traffic demand per cell exceeds some value (about 2 Mbits/s in our numerical example), the gain induced by opportunism compensates the degradation induced by fading compared to AWGN. Partial results were presented at ComNet in 2009 [62].

6.1.1.2. Impact of Shadowing on QoS in Cellular Networks

Shadowing is believed to degrade the quality of service (QoS) in wireless cellular networks. Assuming log-normal shadowing, and studying mobile’s path-loss with respect to the strongest (serving) base station (BS) and the corresponding interference factor (the ratio of the sum of the path-gains form interfering BS’s to the path-gain from the serving BS), which are two key ingredients of the analysis and design of the cellular networks, in [48] we discovered a more subtle reality. We observe, as commonly expected, that a strong variance of the shadowing increases the mean path-loss with respect to the serving BS, which in consequence, may compromise QoS. However, in some cases, an increase of the variance of the shadowing can significantly reduce the mean interference factor and, in consequence, improve some QoS metrics in interference limited systems, provided the handover policy selects the strongest BS as the serving one. We exemplify this phenomenon, similar to stochastic resonance, studying the blocking probability in regular, hexagonal networks in a semi-analytic manner, using a spatial version of the Erlang’s loss formula combined with Kaufman-Roberts algorithm. More detailed probabilistic analysis explains that increasing variance of the log-normal shadowing amplifies the ratio between the strongest signal and all other signals thus reducing the interference. The above observations might shed new light, in particular on the design of indoor communication scenarios. Partial results were presented at IFIP WMNC’2010 [63].
6.1.1.3. Self-Optimization of Radio Resources in Cellular Networks

In [65], we developed mathematical and algorithmic tools for the self-optimization of mobile cellular networks. Scalable algorithms which are based on local measurements and do not require heavy coordination among the wireless devices were proposed. We focused on the optimization of transmit power and of user association. The method is applicable to both joint and separate optimizations. The global utility minimized is linked to potential delay fairness. The distributed algorithm adaptively updates the system parameters and achieves global optimality by measuring SINR and interference. The algorithms are built on Gibbs’ sampler and offer a unified framework that can be easily used for different purposes.

In [32], we investigated the joint optimization of radio resources in heterogeneous cellular networks made of a juxtaposition of macro and small cells. We showed that within this context, it is essential to use algorithms able to simultaneously solve the problems of channel selection, user association and power control. In such networks, the unpredictability of the cell and user patterns also requires self-optimized schemes. We proposed a generalized solution which is based on Gibbs’ sampler. It can be implemented in a distributed way and nevertheless achieves minimal system-wide potential delay. Results show that it is effective in both throughput and energy efficiency.

In [35], we extended it to an autonomous radio resource allocation and optimization scheme that chooses the transmit power and precoding vector among codebooks for multiple antennas transmitters to improve spectral and power efficiency and provide user fairness. Network self-optimization is an essential feature for supporting the cell densification in future wireless cellular systems. Besides, we included power pricing to parametrize and to enhance the energy efficiency. Simulation results show that the proposed scheme can outperform today’s default modes of operation in network throughput, energy efficiency, and user fairness.

Three patents were filed under the INRIA/Alcatel–Lucent joint laboratory.

6.1.1.4. Best Signal Quality in a Wireless Network

In a wireless network composed of randomly scattered nodes, the characterization of the distribution of the best signal quality received from a group of nodes is of primary importance for many network design problems. The thesis of Van Minh Nguyen [7] developed a framework for analyzing this distributions using shot noise models for the interference field. The joint distribution of the interference and the maximum signal strength was identified. The best signal quality can be represented as a function of these two quantities. Particular practical scenarios were also analyzed where explicit expressions can be obtained.

6.1.1.5. Cellular Network Tomography

The Foschini-Miljanic’s [67] algorithm is used for power control in cellular networks when users require a fixed bit rate. It leads to an optimal choice of power by the users in a distributed way when such a solution exists. If the users are too greedy or too many, the network saturates, and it is not possible to provide the required bit rates. We have been working on the question of residual bandwidth estimation in [61]. The residual bandwidth of a user is defined as the rate that this user should have to saturate the network when all other users stick to their initial rate requirement and all users use power control. The aim is to determine the residual bandwidth of a given user by local measurements. We showed that by simply changing their SINR target slightly and by listening to the evolution of interference, users can locally inverse Foschini-Miljanic’s algorithm and compute their residual bandwidth.

6.1.1.6. Coverage in Cellular Networks

Cellular networks are usually modeled by placing the base stations according to a regular geometry such as a grid, with the mobile users scattered around the network either as a Poisson point process (i.e. uniform distribution) or deterministically. These models have been used extensively for cellular design and analysis but suffer from being both highly idealized and not very tractable. Thus, complex simulations are used to evaluate key metrics such as coverage probability for a specified target rate (equivalently, the outage probability) or average/sum rate. More tractable models have long been desirable. In a joint work with J. Andrews and R. Ganti [UT Austin, USA] [9] and [34], we developed general models for multi-cell signal-to-noise-plus-interference ratio (SINR) based on homogeneous Poisson point processes and derived the coverage probability
and rate. Under very general assumptions, the resulting expressions for the SINR cumulative distribution function involve quickly computable integrals, and in some important special cases of practical interest these integrals can be simplified to common integrals (e.g., the Q-function) or even to exact and quite simple closed-form expressions. We also derived the mean rate, and then the coverage gain (and mean rate loss) from static frequency reuse. We compared the coverage predictions obtained by this approach to the standard grid model and an actual base station deployment. We observed that the proposed model is pessimistic (a lower bound on coverage) whereas the grid model is optimistic. In addition to being more tractable, the proposed model may better capture the increasingly opportunistic and dense placement of base stations in urban cellular networks with highly variable coverage radii.

Cellular networks are in a major transition from a carefully planned set of large tower-mounted base-stations (BSs) to an irregular deployment of heterogeneous infrastructure elements that often additionally includes micro, pico, and femtocells, as well as distributed antennas. In a collaboration with H. Dhillon, J. Andrews and R. Ganti [UT Austin, USA] [66], we extended the approach of we developed a model for a downlink heterogeneous cellular network (HCN) consisting of K tiers of randomly located BSs, where each tier may differ in terms of average transmit power, supported data rate and BS density. Assuming a mobile user connects to the strongest candidate BS, the resulting Signal-to-Interference-plus-Noise-Ratio (SINR) is greater than 1 when in coverage, Rayleigh fading, we derived an expression for the probability of coverage (equivalently outage) over the entire network under both open and closed access. One interesting observation for interference-limited open access networks is that at a given SINR, adding more tiers and/or BSs neither increases nor decreases the probability of coverage or outage when all the tiers have the same SINR threshold.

6.1.2. Mobile Ad Hoc Networks

A MANET is made of mobile nodes which are at the same time terminals and routers, connected by wireless links, the union of which forms an arbitrary topology. The nodes are free to move randomly and organize themselves arbitrarily. Important issues in such a scenario are connectivity, medium access (MAC), routing and stability. This year, we worked on the analysis of MAC and routing protocols in multi-hop MANETS in collaboration with Paul Mühlethaler [INRIA HIPERCOM], and on a game theoretic view of Spatial Aloha in collaboration with E. Altman and M.K. Hanawal [INRIA MAESTRO] [68].

6.1.2.1. Improvement of CSMA/CA's Spatial Reuse

The most popular medium access mechanism for such ad hoc networks is CSMA/CA with RTS/CTS. In CSMA-like mechanisms, spatial reuse is achieved by implementing energy based guard zones. In a new collaboration with Qualcomm ([26] and [14]), we considered the problem of simultaneously scheduling the maximum number of links that can achieve a given signal to interference ratio (SIR). Using tools from stochastic geometry, we studied and maximized the medium access probability of a typical link. Our contributions are two-fold: (i) We showed that a simple modification to the RTS/CTS mechanism, viz., changing the receiver yield decision from an energy-level guard zone to an SIR guard zone, leads to performance gains; and (ii) We showed that this combined with a simple modification to the transmit power level – setting it to be inversely proportional to the square root of the link gain – leads to significant improvements in network throughput. Further, this simple power-level choice is no worse than a factor of two away from optimal over the class of all "local" power level selection strategies for fading channels, and further is optimal in the non-fading case. The analysis relies on an extension of the Matérn hard core point process which allows us to quantify both these SIR guard zones and this power control mechanism.

6.1.2.2. Opportunistic versions of CSMA/CA

In collaboration with Gustavo de Veciana and Yuchul Kim [UT Austin, ECE] we studied the benefits of channel-aware (opportunistic) scheduling of transmissions in ad-hoc networks using CSMA/CA [36]. The key challenge in optimizing the performance of such systems is finding a good compromise among three interdependent quantities, the density and channel quality of the scheduled transmitters, and the resulting interference at receivers. We propose two new channel-aware slotted CSMA protocols: opportunistic CSMA (O-CSMA) and quantile-based CSMA (QT-CSMA) and develop stochastic geometric models allowing us to quantify their performance in terms of spatial reuse and spatial fairness. When properly optimized these
protocols offer substantial improvements in terms of both of these metrics relative to CSMA - particularly when the density of nodes is moderate to high. Moreover, we show that a simple version of QT-CSMA can achieve robust performance gains without requiring careful parameter optimization. The paper supports the case that the benefits associated with channel-aware scheduling in ad hoc networks, as in centralized base station scenarios, might far outweigh the associated overhead, and this can be done robustly using a QT-CSMA like protocol.

6.1.3. Cognitive Radio Networks

We wrote a survey [22] on the probabilistic framework which can be used to model and analyze cognitive radio networks using various classes of MAC protocols (including carrier sensing based multiple access schemes and Aloha schemes). For each model, analytical results were derived for important performance metrics. This leads to a quantification of the interplay between primary and secondary users in such networks.

6.1.4. Generic Wireless Networks

6.1.4.1. Power Control in Wireless Networks

In [10], in collaboration with N. Bambos, [Stanford] and N. Gast [EPFL], we formulated a delay-power control (DPC) scheme for wireless networking, which balances delay against transmitter power on each wireless link. The DPC scheme is scalable, as each link autonomously updates its power based on the interference observed at its receiver; no cross-link communication is required. It is shown that DPC converges to a unique equilibrium power and several ke properties are established, concerning the nature of channel bandwidth sharing achieved by the links. The DPC scheme is contrasted to the well-known Foschini-Miljanic (FM) formulation for transmitter power control in wireless networks, and some key advantages are established. Based on the DPC and FM schemes, two protocols are developed, which leverage adaptive tuning of DPC parameters. One of them is inspired by TCP and exhibits analogous behavior.

In [21], we studied the weighted sum rate maximization problem in wireless networks consisting of multiple source-destination pairs. The optimization problem is to maximize a weighted sum of data rates by adjusting the power of each user. The problem is in general a non-convex optimization problem that will lead to multiple local maxima. A Gauss-Seidel type iterative power control algorithm was presented. We showed by simulation that the proposed algorithm converges to the global maximum with very high probability, if we initialize the initial power allocation uniformly at random. The proposed algorithm also has the favorable properties that only simple operations are needed in each iteration, and the convergence is fast. Performance comparison under different user densities has also indicated its effectiveness. Finally, we discussed some simple and optimal power allocation strategies under special cases of the problem if the network can be represented by a certain approximation.

6.1.4.2. Simultaneous Decoding

In [15], in collaboration with A. El Gamal [Stanford, USA] and D. Tse [UC Berkeley, USA], we analyzed a network made of a collection of transmitter-receiver links where each link is considered to be part of a Multiple Access Channel (MAC) together with a collection of co-transmitters, rather than treating the messages of the latter as noise. This MAC extension is meant to improve the rate of the link and not to decode the messages of the co-transmitters. The necessary and sufficient condition for the feasibility of some rate when using successive interference cancellation and simultaneous decoding were provided. The reasons why simultaneous decoding is preferable to successive interference cancellation were also given. The gain obtained when using this type of simultaneous decoding rather than treating interference as noise was then quantified in a network made of a large random collection of such links. The gains in coverage and in rate were analyzed in terms of ensemble averages, evaluated using stochastic geometry. Closed form or integral expressions were obtained for the outage/coverage probability in networks where nodes are randomly distributed like a Poisson point process on an infinite plane. In the CDMA limit (large bandwidth, low SINR per hertz, high density), the ensemble average of the link rates tends to 0 when interference is treated as noise whereas it tends to a positive constant when simultaneous decoding of infinite order is used. The whole analysis was conducted in the AWGN case.
6.2. Network Dynamics


This traditional research topic of TREC has several new threads like perfect simulation, active probing or Markov decision.

6.2.1. Network Calculus

Network calculus is a theory that aims at computing deterministic performance guarantees in communication networks. This theory is based on the (min,plus) algebra. Flows are modeled by an arrival curve that upper-bounds the amount of data that can arrive during any interval, and network elements are modeled by a service curve that gives a lower bound on the amount of service offered to the flows crossing that element. Worst-case performances are then derived by combining these curves.

6.2.1.1. Performance bounds networks with static priorities

In cooperation with Aurore Junier [INRIA/IRISA], we present in [29] algorithms to compute worst-case performance upper bounds when the service policy is static priorities, using linear programming. Linear programming does not lead to tight bounds, but when combining this method with (min,plus) methods, we obtain bounds that outperform the already known bounds. Also, we prove that in tandem networks, the the worst-case performance bound under arbitrary multiplexing can be obtain by a policy with static priorities, the “shortest-destination first” policy.

6.2.1.2. Feed-forward networks with wormhole routing discipline

In collaboration with Bruno Gaujal [INRIA Rhone Alpes] and Nadir Farhi [IFFSTAR] we are working on a model of performance bound calculus on feed-forward networks where data packets are routed under wormhole routing discipline. We are interested in determining maximum end-to-end delays and backlogs for packets going from a source node to a destination node, through a given virtual path in the network. Our objective is to give a “network calculus” approach to calculate the performance bounds. For this, we propose a new concept of curves that we call packet curves. The curves permit to model constraints on packet lengths for data flows, when the lengths are allowed to be different. We used this new concept to propose an approach for calculating residual services for data flows served under non preemptive service disciplines. This notion also enabled us to differentiate different classes of service policies: those that are based on a packet count (like round-robin and its generalized version), where the packet curve will be useful to tighten the bounds computed, and those that are based on the amount of data served (FIFO, priorities), where it won’t be useful. These results can be found in [44] and have been presented in ILAS 2011.

6.2.1.3. Composition of service curves in Network Calculus

In envelope-based models for worst-case performance evaluation like Network Calculus or Real-Time Calculus, several types of service curves have been introduced to quantify some deterministic service guarantees. We compare those different classes of service curves regarding the composition (servers in tandem) and individual service curves (when several flows share a server, what service curve can be guaranteed to each of the flows?). In short, there are two main classes of service curves, simple and strict service curves. Individual service curve can not always be computed when simple service curves are considered, and strict service curves is not a stable class regarding the two operations described. We show that there can be no equivalence between the two main classes of service curves and that no notion of service curve in-between can be defined, that behaves well for the composition. We complete this study by studying other classes of service curves from this viewpoint. These results have been presented in [28]

6.2.1.4. Residuation in (max,plus) automata

With Éric Badouel, Philippe Darondeau [INRIA/IRISA] and Jan Komenda [Institute of Mathematics, Brno], we study in [27] the decidability of existence and the rationality of delay controllers for systems with time weights in the tropical and interval semirings. Depending on the (max,+) or (min,+) rationality of the series specifying the controlled system and the control objective, cases are identified where the controller series
defined by residuation is rational, and when it is positive (i.e., when delay control is feasible). When the control objective is specified by a tolerance, i.e. by two bounding rational series, a nice case is identified in which the controller series is of the same rational type as the system specification series.

6.2.2. Queueing Theory and Active Probing

6.2.2.1. Inverse Problems

Active probing began by measuring end-to-end path metrics, such as delay and loss, in a direct measurement process which did not require inference of internal network parameters. The field has since progressed to measuring network metrics, from link capacities to available bandwidth and cross traffic itself, which reach deeper and deeper into the network and require increasingly complex inversion methodologies. The thesis of B. Kauffmann [6] investigates this line of thought as a set of inverse problems in queueing theory. Queueing theory is typically concerned with the solution of direct problems, where the trajectory of the queueing system, and laws thereof, are derived based on a complete specification of the system, its inputs and initial conditions. Inverse problems aim to deduce unknown parameters of the system based on partially observed trajectories. A general definition of the inverse problems in this class was provided and the key variants were mapped out: the analytical methods, the statistical methods and the design of experiments. We also show how this inverse problem viewpoint translates to the design of concrete Internet probing applications.

Inverse problems in bandwidth sharing networks theory were also investigated. A bandwidth sharing networks allocates the bandwidth to each flow in order to maximize a given utility function (typically an $\alpha$-fairness), with the constraints given by the capacity of the different servers. In particular, it has been shown that the equilibrium distribution of the bandwidth allocated by TCP to many competing connections is oscillating around an $\alpha$-fair allocation. As such, the theory of bandwidth sharing network is a high-level viewpoint of networks. The meaning of inverse problems in this theory, and their relation to the active probing paradigm are analyzed. In two simple examples of network, the capacity of the different servers and the flow population can estimated, and an algorithm to perform this estimation was proposed.

6.2.2.2. Internet Tomography

Most active probing techniques suffer of the “Bottleneck” limitation: all characteristics of the path after the bottleneck link are erased and unreachable. we are currently investigating a new tomography technique, based on the measurement of the fluctuations of point-to-point end-to-end delays, and allowing one to get insight on the residual available bandwidth along the whole path. For this, we combined classical queueing theory models with statistical analysis to obtain estimators of residual bandwidth on all links of the path. These estimators were proved to be tractable, consistent and efficient. In [59] we evaluated their performance with simulation and trace-based experiments.

Lately this method has been generalized in [72] to a probing multicast tree instead of a single path. This work deals with the complexity of the combinatorials in trees, and gives an explicit formula for the iteration of the Expectation-Maximization (E-M) algorithm. The E-M algorithm is notoriously slow, and we provided three speed-up techniques which are effective in our case (up to a factor $10^3$ in the computation time). These techniques are general, and can be applied to other instances of E-M, or even several other iterative algorithms.

6.2.3. Perfect Sampling of Queueing Systems

Propp and Wilson introduced in 1996 a perfect sampling algorithm that uses coupling arguments to give an unbiased sample from the stationary distribution of a Markov chain on a finite state space $\mathcal{X}$. In the general case, the algorithm starts trajectories from all $x \in \mathcal{X}$ at some time in the past until time $t = 0$. If the final state is the same for all trajectories, then the chain has coupled and the final state has the stationary distribution of the Markov chain. Otherwise, the simulations are started further in the past. This technique is very efficient if all the events in the system have appropriate monotonicity properties. However, in the general (non-monotone) case, this technique requires that one consider the whole state space, which limits its application only to chains with a state space of small cardinality.
6.2.3.1. Piecewise Homogeneous Events

In collaboration with Bruno Gaujal [INRIA Grenoble - Rhone-Alpes], we proposed in [47] a new approach for the general case that only needs to consider two trajectories. Instead of the original chain, we used two bounding processes (envelopes) and we showed that, whenever they couple, one obtains a sample under the stationary distribution of the original chain. We showed that this new approach is particularly effective when the state space can be partitioned into pieces where envelopes can be easily computed. We further showed that most Markovian queueing networks have this property and we propose efficient algorithms for some of them.

The envelope technique has been implemented in a software tool PSI2 (see Section 5.2).

6.2.3.2. Acceleration of Perfect Sampling by Skipping Events

In collaboration with Bruno Gaujal [INRIA Grenoble - Rhone-Alpes], we proposed a new method to speed up perfect sampling of Markov chains by skipping passive events during the simulation [38]. We showed that this can be done without altering the distribution of the samples. This technique is particularly efficient for the simulation of Markov chains with different time scales such as queueing networks where certain servers are much faster than others. In such cases, the coupling time of the Markov chain can be arbitrarily large while the runtime of the skipping algorithm remains bounded. This was further illustrated by several experiments that also show the role played by the entropy of the system in the performance of our algorithm.

6.2.3.3. Aggregated Envelopes

When the cardinality of the state space is so huge that even storing the state of the Markov chain becomes challenging, we propose to combine the ideas of bounding processes and the aggregation of Markov chains [30]. We illustrate the proposed approach of aggregated envelope bounding chains on queueing models with joint arrivals and joint services, often referred to in the literature as assemble-to-order systems. Due to the finite capacity, and coupling in arrivals and services, the exact solving techniques are inefficient for larger problem instances. For instance, for the service tools model proposed by Vliegen and Van Houtum (2009), the aggregated envelope method reduces exponentially the dimension of the state space and allows effective perfect sampling algorithms. We also provide bounds for the coupling time, under the high service rate assumptions.

6.2.4. Markov Chains and Markov Decision Processes

Solving Markov chains is in general difficult if the state space of the chain is very large (or infinite) and lacking a simple repeating structure. One alternative to solving such chains is to construct models that are simple to analyze and provide bounds for a reward function of interest. The bounds can be established by using different qualitative properties, such as stochastic monotonicity, convexity, submodularity, etc. In the case of Markov decision processes, similar properties can be used to show that the optimal policy has some desired structure (e.g. the critical level policies).

6.2.4.1. Stochastic Monotonicity

In collaboration with Jean-Michel Fourneau [PRiSM, Université de Versailles Saint-Quentin] we consider two different applications of stochastic monotonicity in performance evaluation of networks [18]. In the first case, we assume that a Markov chain of the model depends on a parameter that can be estimated only up to a certain level and we have only an interval that contains the exact value of the parameter. Instead of taking an approximated value for the unknown parameter, we show how we can use the monotonicity properties of the Markov chain to take into account the error bound from the measurements. In the second application, we consider a well known approximation method: the decomposition into submodels. In such an approach, models of complex networks are decomposed into submodels whose results are then used as parameters for the next submodel in an iterative computation. One obtains a fixed point system which is solved numerically. In general, we have neither an existence proof of the solution of the fixed point system nor a convergence proof of the iterative algorithm. Here we show how stochastic monotonicity can be used to answer these questions. Furthermore, monotonicity properties can also help to derive more efficient algorithms to solve fixed point systems.
6.2.4.2. Componentwise Bounds

In collaboration with Jean-Michel Fourneau [PRiSM, Université de Versailles Saint-Quentin] we proposed an iterative algorithm to compute component-wise bounds of the steady-state distribution of an irreducible and aperiodic Markov chain [17]. These bounds are based on very simple properties of \((\max, +)\) and \((\min, +)\) sequences. We showed that, under some assumptions on the Markov chain, these bounds converge to the exact solution. In that case we have a clear tradeoff between computation and the tightness of bounds. Furthermore, at each step we know that the exact solution is within an interval, which provides a more effective convergence test than usual iterative methods.

6.2.4.3. Markov Reward Processes and Aggregation

In a joint work with I.M. H. Vliegen [University of Twente, The Netherlands] and A. Scheller-Wolf [Carnegie Mellon University, USA] [19], we presented a new bounding method for Markov chains inspired by Markov reward theory: Our method constructs bounds by redirecting selected sets of transitions, facilitating an intuitive interpretation of the modifications of the original system. We show that our method is compatible with strong aggregation of Markov chains; thus we can obtain bounds for an initial chain by analyzing a much smaller chain. We illustrated our method by using it to prove monotonicity results and bounds for assemble-to-order systems.

6.2.4.4. Critical Level Policies in Controlled Queuing Systems

In a joint work with Emmanuel Hyon [University of Paris Ouest Nanterre La Defense and LIP6] [39], we consider a single-item lost sales inventory model with different classes of customers. Each customer class may have different lost sale penalty costs. We assume that the demands follow a Poisson process and we consider a single replenishment hypoexponential server. We give a Markov decision process associated with this optimal control problem and prove some structural properties of its dynamic programming operator. This allows us to show that the optimal policy is a critical level policy. We also discuss some possible extensions to other replenishment distributions and give some numerical results for the hypoexponential server case.

6.2.5. Dynamic Systems with Local Interactions

Dynamic systems with local interactions can be used to model problems in distributed computing: gathering a global information by exchanging only local information. The challenge is two-fold: first, it is impossible to centralize the information (cells are indistinguishable); second, the cells contain only a limited information (represented by a finite alphabet \(A; A = \{0, 1\}\) in our case). Two natural instantiations of dynamical systems are considered, one with synchronous updates of the cells, and one with asynchronous updates. In the first case, time is discrete, all cells are updated at each time step, and the model is known as a Probabilistic Cellular Automaton (PCA) (e.g. Dobrushin, R., Kryukov, V., Toom, A.: Stochastic cellular systems: ergodicity, memory, morphogenesis, 1990). In the second case, time is continuous, cells are updated at random instants, at most one cell is updated at any given time, and the model is known as a (finite range) Interacting Particle System (IPS) (e.g. Liggett, T.M.: Interacting particle systems, 2005).

6.2.5.1. Density Classification on Infinite Lattices and Trees

In a joint work with N. Fatès [INRIA Nancy – Grand-Est], J. Mairesse and I. Marcovici [LIAFA, CNRS and Université Paris 7] [46] we consider an infinite graph with nodes initially labeled by independent Bernoulli random variables of parameter \(p\). We address the density classification problem, that is, we want to design a (probabilistic or deterministic) cellular automaton or a finite-range interacting particle system that evolves on this graph and decides whether \(p\) is smaller or larger than \(1/2\). Precisely, the trajectories should converge (weakly) to the uniform configuration with only \(0's\) if \(p < 1/2\), and only \(1's\) if \(p > 1/2\). We present solutions to that problem on \(\mathbb{Z}^d\), for any \(d \geq 2\), and on the regular infinite trees. For \(\mathbb{Z}\), we propose some candidates that we back up with numerical simulations.

6.2.6. Stochastic Stability

6.2.6.1. Ergodicity of Probabilistic Cellular Automata
In a joint work with J. Mairesse and I. Marcovici [LIAFA, CNRS and Université Paris 7] [31], we considered ergodicity properties of probabilistic cellular automata (PCA). A classical cellular automaton (CA) is a particular case of PCA. For a 1-dimensional CA, we proved that ergodicity is equivalent to nilpotency, and is therefore undecidable. We then proposed an efficient perfect sampling algorithm for the invariant measure of an ergodic PCA. Our algorithm does not assume any monotonicity properties of the local rule. It is based on a bounding process which is shown to be also a PCA. We then focused on the PCA Majority, whose asymptotic behavior is unknown, and performed numerical experiments using the perfect sampling procedure.

6.2.6.2. Spatial Queues

In a joint work with S. Foss [Heriot–Watt University, UK] [13], we considered a queue where the server is the Euclidean space, the customers are random closed sets of the Euclidean space arriving according to a Poisson rain and where the discipline is a hard exclusion rule: no two intersecting random closed sets can be served at the same time. We use the max plus algebra and Lyapunov exponents to show that under first come first serve assumptions, this queue is stable for a sufficiently small arrival intensity. We also discuss the percolation properties of the stationary regime of the random closed sets in the queue.

6.3. Economics of Networks

Participants: François Baccelli, Emilie Coupechoux, Marc Lelarge.

6.3.1. Diffusion and Cascading Behavior in Random Networks

The spread of new ideas, behaviors or technologies has been extensively studied using epidemic models. In [69], we considered a model of diffusion where the individuals’ behavior is the result of a strategic choice. We studied a simple coordination game with binary choice and give a condition for a new action to become widespread in a random network. We also analyze the possible equilibria of this game and identify conditions for the coexistence of both strategies in large connected sets. Finally we look at how can firms use social networks to promote their goals with limited information.

Our results differ strongly from the one derived with epidemic models. In particular, we showed that connectivity plays an ambiguous role: while it allows the diffusion to spread, when the network is highly connected, the diffusion is also limited by high-degree nodes which are very stable. In the case of a sparse random network of interacting agents, we computed the contagion threshold for a general diffusion model and showed the existence of (continuous and discontinuous) phase transitions. We also computed the minimal size of a seed of new adopters in order to trigger a global cascade if these new adopters can only be sampled without any information on the graph. We showed that this minimal size has a non-trivial behavior as a function of the connectivity. Our analysis extends methods developed in the random graphs literature based on the properties of empirical distributions of independent random variables, and leads to simple proofs.

6.3.2. Impact of Clustering on Diffusions and Contagions in Random Networks

In [33] we extend some results of the previous results to a model of random graphs having both a given degree distribution and a tunable clustering coefficient. This work shed new light on the impact of clustering on the spread of new ideas, technologies, viruses or worms. We consider two types of growth processes: the (classical SI) diffusion model, and the contagion model, which is inspired by a simple coordination game played on the network and is characterized by a threshold rule and a random seed. While clustering inhibits the diffusion process (on regular graphs), 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.

6.3.3. Economic Value of User Localization in Wireless Networks

The defining characteristic of wireless and mobile networking is user mobility, and related to it is the ability for the network to capture (at least partial) information on where users are located and how users change location over time. Information about location is becoming critical, and therefore valuable, for an increasingly larger number of location-based or location-aware services. A key open question, however, is how valuable exactly this information is. Our goal in this paper is to help understand and estimate the economics, or the value of location information.
In a joint work with J. Bolot [Sprint ATL, USA], [25], we addressed in particular the value of different granularities of location information, for example how much more valuable is it to know the GPS location of a mobile user compared to only knowing the access point, or the cell tower, that the user is associated with. We made three main contributions. First, we presented novel models, which capture the location-based economic activity of mobile users. Second, we derived closed-form analytic solutions for the economic value generated by those users. Third, we augmented the models to consider uncertainty about the users’ location, and derived expressions for the economic value generated with different granularities of location information.

6.4. Point Processes, Stochastic Geometry and Random Geometric Graphs

Participants: François Baccelli, Bartłomiej Blaszczyszyn, Pierre Brémaud, Yogeshwaran Dhandapani, Kumar Gaurav, Mir Omid Haji Mirsadeghi, Justin Salez.

6.4.1. Comparison of Clustering and Percolation of Point Processes and Random Graphs

Heuristics indicate that point processes exhibiting clustering of points have larger critical radius \( r_c \) for the percolation of their continuum percolation models than spatially homogeneous point processes. It has already been shown in [64], [73] that the directionally convex (\( dcx \)) ordering of point processes is suitable to compare their clustering tendencies. Hence, it was tempting to conjecture that \( r_c \) is increasing in \( dcx \) order. Some numerical evidences support this conjecture for a special class of point processes, called perturbed lattices, which are "toy models" for determinantal and permanental point processes. However the conjecture is not true in full generality. In 2011 we have prepared three publications on this subject.

6.4.1.1. On comparison of clustering properties of point processes

In [52] we provide a large class of perturbed lattice point processes, monotone in \( dcx \) order and comparable to Poisson point processes that is commonly considered as the reference model in the comparative study of clustering phenomena. We also introduce a weaker order based on the comparison of only 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 weaker sense to the Poisson point process of the same mean measure.

6.4.1.2. Clustering and percolation of point processes

In [49] we show that simple, stationary point processes of a given intensity on \( \mathbb{R}^d \), having void probabilities and factorial moment measures smaller than those of a homogeneous Poisson point process of the same intensity, admit uniformly non-degenerate lower and upper bounds on the critical radius \( r_c \) for the percolation of their continuum percolation models. Examples are negatively associated point processes and, more specifically, determinantal point processes. More generally, we show that point processes \( dcx \) smaller than a homogeneous Poisson point processes (for example perturbed lattices) exhibit phase transitions in certain percolation models based on the level-sets of additive shot-noise fields of these point processes. Examples of such models are \( k \)-percolation and SINR-percolation models. We also construct a Cox point process with degenerate critical radius \( r_c = 0 \), that is \( dcx \) larger than a given homogeneous Poisson point process. This is a counterexample for the aforementioned conjecture in the full generality.

6.4.1.3. Ordering of non-standard critical radii

As explained above, heuristically one expects finiteness of the critical radii for percolation of sub-Poisson point processes. However, in [49] we have show that it is non-zero as well. In a more elaborate paper [50] we present a reasoning as to why this non-triviality is to be expected. Specifically, we defined two (nonstandard) critical radii for percolation of the Boolean model, called the lower and upper critical radii, and related, respectively, to the finiteness of the expected number of void circuits around the origin and asymptotic of the expected number of long occupied paths from the origin in suitable discrete approximations of the continuum model. These radii sandwich the usual critical radius \( r_c \) for percolation of the Boolean model. We show that \( dcx \) order preserves the upper critical radii and reverses the lower critical radii.
6.4.1.4. Local weak convergence and stochastic comparison

Many random models are parametrized by the size of the model, and the essential properties of the model are the asymptotic ones as the size of the graph tends to infinity. In the master thesis [57] we show that the theory of local weak converge provides a natural setting to investigate stochastic (convex) ordering of such models. We consider both the geometric context of [71] and the discrete one of Galton-Watson branching process and Configuration Model, cf [5]. In this latter case we define and study a convex order in the context of random trees and graphs which converge in the local weak sense. In particular, we’re interested in the effect of ordering on percolation. It turns out that while in the case of Galton-Watson trees, convex ordering leads to the ordering of percolation probabilities, we cannot conclude this in the case of configuration model. In this case, we could only obtain the ordering of percolation thresholds.

6.4.1.5. AB random geometric graphs

We investigated percolation in the AB Poisson-Boolean model in d-dimensional Euclidean space, and asymptotic properties of AB random geometric graphs on Poisson points in [0, 1]^d. The AB random geometric graph we studied is a generalization to the continuum of a bi-partite graph called the AB percolation model on discrete lattices. Such an extension is motivated by applications to secure communication networks and frequency division duplex networks. The AB Poisson Boolean model is defined as a bi-partite graph on two independent Poisson point processes of intensities λ and μ in the d-dimensional Euclidean space in the same manner as the usual Boolean model with a radius r. We showed existence of AB percolation for all d ≥ 2, and derived bounds for a critical intensity. Further, in d = 2, we characterize a critical intensity. The set-up for AB random geometric graphs is to construct a bi-partite graph on two independent Poisson point process of intensities n and cn in the unit cube. We provided almost sure asymptotic bounds for the connectivity threshold for all c > 0 and a suitable choice of radius cut-off functions r_n(c). Further for c < c_0, we derived a weak law result for the largest nearest neighbor radius. This work, which was a part of the PhD thesis [73] will appear in [23].

6.4.2. Random Packing Models

Random packing models (RPM) are point processes (p.p.s) where points which “contend” with each other cannot be simultaneously present. These p.p.s play an important role in many studies in physics, chemistry, material science, forestry and geology. For example, in microscopic physics, chemistry and material science, RPMs can be used to describe systems with hard-core interactions. Applications of this type range from reactions on polymer chains, chemisorption on a single-crystal surface, to absorption in colloidal systems. In these models, each point (molecule, particle, ...) in the system occupies some space, and two points with overlapping occupied space contend with each other. Another example is the study of seismic and forestry data patterns, where RPMs are used as a reference model for the data set under consideration. In wireless communications, RPMs can be used to model the users simultaneously accessing the medium in a wireless network using Carrier Sensing Medium Access (CSMA). In this context, each point (node, user, transmitter, ...) does not occupy space but instead generates interference to other points in the network. Two points contend with each other if either of them generates too much interference to the other. Motivated by this kind of application, we studied in [70] the generating functionals of several models of random packing processes: the classical Matérn hard-core model; its extensions, the k-Matérn models and the ∞-Matérn model, which is an example of random sequential packing process. The main new results are: 1) A sufficient condition for the ∞-Matérn model to be well-defined (unlike the other two, the ∞-Matérn model may not be well-defined on unbounded space); 2) the generating functional of the resulting point process which is given for each of the three models as the solution of a differential equation; 3) series representation and bounds on the generating functional of the packing models; 4) moment measures and other useful properties of the considered packing models which are derived from their generating functionals.

6.4.3. Extremal and Additive Matérn Point Processes

In the simplest Matérn point processes, one retains certain points of a Poisson point process in such a way that no pairs of points are at distance less than a threshold. This condition can be reinterpreted as a threshold condition on an extremal shot–noise field associated with the Poisson point process. In a joint work with P.
Bermolen (Universidad de la República, Montevideo, Uruguay) [11], we studied extensions of Matérn point processes where one retains points that satisfy a threshold condition based on an additive shot–noise field of the Poisson point process. We provide an analytical characterization of the intensity of this class of point processes and we compare the packing obtained by the extremal and additive schemes and certain combinations thereof.

6.4.4. Spatial Birth and Death Point Processes

In collaboration with F. Mathieu [INRIA GANG] and Ilkka Norros [VTT, Finland], we started studying a new spatial birth and death point process model where the death rate is a shot noise of the point configuration [60]. We showed that the spatial point process describing the steady state exhibits repulsion. We studied two asymptotic regimes: the fluid regime and the hard–core regime. We derived closed form expressions for the mean (and in some cases the law) of the latency of points as well as for the spatial density of points in the steady state of each regime.

6.4.5. Information Theory and Stochastic Geometry

In a joint work with V. Anantharam [UC Berkeley], [58], we studied the Shannon regime for the random displacement of stationary point processes. Let each point of some initial stationary point process in \( n \)-dimensional Euclidean space give rise to one daughter point, the location of which is obtained by adding a random vector to the coordinates of the mother point, with all displacement vectors independently and identically distributed for all points. The decoding problem is then the following one: the whole mother point process is known as well as the coordinates of some daughter point; the displacements are only known through their law; can one find the mother of this daughter point? The Shannon regime is that where the dimension \( n \) tends to infinity and where the logarithm of the intensity of the point process is proportional to \( n \). We showed that this problem exhibits a sharp threshold: if the sum of the proportionality factor and of the differential entropy rate of the noise is positive, then the probability of finding the right mother point tends to 0 with \( n \) for all point processes and decoding strategies. If this sum is negative, there exist mother point processes, for instance Poisson, and decoding strategies, for instance maximum likelihood, for which the probability of finding the right mother tends to 1 with \( n \). We then used large deviations theory to show that in the latter case, if the entropy spectrum of the noise satisfies a large deviation principle, then the error probability goes exponentially fast to 0 with an exponent that is given in closed form in terms of the rate function of the noise entropy spectrum. This was done for two classes of mother point processes: Poisson and Matérn. The practical interest to information theory comes from the explicit connection that we also establish between this problem and the estimation of error exponents in Shannon’s additive noise channel with power constraints on the codewords.

We currently investigate extensions of this approach to network information theoretic channels.

6.4.6. Navigation on Point Processes and Graphs

In [12], we studied optimal navigations in wireless networks in terms of first passage percolation on some space–time SINR graph. We established both “positive” and “negative” results on the associated the percolation delay rate (delay per unit of Euclidean distance, also called time constant in the classical terminology of percolation). The latter determines the asymptotics of the minimum delay required by a packet to progress from a source node to a destination node when the Euclidean distance between the two tends to infinity. The main negative result states that the percolation delay rate is infinite on the random graph associated with a Poisson point process under natural assumptions on the wireless channels. The main positive result states that when adding a periodic node infrastructure of arbitrarily small intensity to the Poisson point process, the percolation delay rate is positive and finite.

A new direction of research was initiated aiming at defining a new class of measures on a point process which are invariant under the action of a navigation on this point process. This class of measures has properties similar to Palm measures of stationary point processes; but they cannot be defined in the classical framework of Palm measures.

6.5. Random Graphs and Combinatorial Optimization

Participants: Hamed Amini, Emilie Coupechoux, Mathieu Leconte, Marc Lelarge, Justin Salez.
6.5.1. Rank of Large Random Graphs

In [16], with Charles Bordenave [CNRS-Université de Toulouse], we investigated the rank of the adjacency matrix of large diluted random graphs: for a sequence of graphs converging locally to a Galton-Watson tree, we provided an explicit formula for the asymptotic multiplicity of the eigenvalue 0 in terms of the degree generating function. In the first part, we showed that the adjacency operator associated with a Galton-Watson tree is self-adjoint with probability one; we analyzed the associated spectral measure at the root and characterize the distribution of its atomic mass at 0. In the second part, we established a sufficient condition for the expectation of this atomic mass to be precisely the normalized limit of the dimension of the kernel of the adjacency matrices of the sequence of graphs. Our proofs borrow ideas from analysis of algorithms, functional analysis, random matrix theory, and statistical physics.

6.5.2. Matchings in infinite graphs

In [43], we proved that for any sequence of (deterministic or random) graphs converging locally, the corresponding sequence of normalized matching numbers converges, and this limit depends only on the limit of the graph sequence. In the particular case where this limit is a unimodular Galton Watson tree, we were able to compute explicitly the value for the limit of the sequence of (normalized) matching numbers. This leads to an explicit formula that considerably extends the well-known one by Karp and Sipser for Erdős-Rényi random graphs.

We considered a natural family of Gibbs distributions over matchings on a finite graph, parameterized by a single positive number called the temperature. The correlation decay technique can be applied for the analysis of matchings at positive temperature and allowed us to establish the weak convergence of the Gibbs marginal as the underlying graph converges locally. However for the zero temperature problem (i.e. maximum matchings), we showed that there is no correlation decay even in very simple cases. By using a complex temperature and a half-plane property due to Heilmann and Lieb, we were able to let the temperature tend to zero and obtained a limit theorem for the asymptotic size of a maximum matching in the graph sequence.

6.5.3. Counting spanning subgraphs subject to local constraints

In [53], we use negative association and local weak convergence to establish the validity of the cavity method for counting spanning subgraphs subject to local constraints. Specifically, the normalized logarithm of the associated generating polynomial (or partition function) is shown to converge along any sequence of graphs whose random weak limit is a tree, and the limit is directly expressed in terms of the unique solution to a limiting cavity equation. On a Galton-Watson tree, the latter simplifies into a recursive distributional equation which can be solved explicitly. As an illustration, we provide an asymptotic formula for the maximal size of a spanning subgraph with maximal degree $b$ in the Erdős-Rényi model with fixed average degree and diverging size, for any $b \in \mathbb{N}$.

6.5.4. Bipartite graph structures for efficient balancing of heterogeneous loads

With Laurent Massoulié [Technicolor], we extend the results obtained previously on the asymptotic size of maximum matchings in random graphs converging locally to Galton-Watson trees to so-called b-matchings (with non-unitary capacity at vertices as well as constraints on individual edges). Compared to the matching case, this involves studying the convergence of a message passing algorithms which transmits vectors instead of single real numbers. We also look further into an application of these results to large scale distributed content service platforms, such as peer-to-peer video-on-demand systems. In this context, the density of maximum b-matchings corresponds to the maximum fraction of simultaneously satisfiable requests, when the service resources are limited and each server can only handle requests for a predetermined subset of the contents which it has stored in memory. An important design aspect of such systems is the content placement strategy onto the servers depending on the estimated content popularities; the results obtained allow to characterize the efficiency of such placement strategies and the optimal strategies in the limit of large storage capacity at servers are determined.
6.5.5. Flooding in Weighted Random Graphs

In a joint work [24] with Moez Draief [Imperial College London], we studied the impact of the edge weights on distances in diluted random graphs. We interpret these weights as delays, and take them as i.i.d exponential random variables. We analyzed the edge flooding time defined as the minimum time needed to reach all nodes from one uniformly chosen node, and the edge diameter corresponding to the worst case edge flooding time. Under some regularity conditions on the degree sequence of the random graph, we showed that these quantities grow as the logarithm of $n$, when the size of the graph $n$ tends to infinity. We also derived the exact value for the prefactors.

These allowed us to analyze an asynchronous randomized broadcast algorithm for random regular graphs. Our results show that the asynchronous version of the algorithm performs better than its synchronized version: in the large size limit of the graph, it will reach the whole network faster even if the local dynamics are similar on average.
6. New Results

6.1. Advances in symbolic parsing with DyALog/FRMG

**Participant:** Éric Villemonte de La Clergerie.

Within the team is developed a wide-coverage French meta-grammar (FRMG) and an efficient hybrid TAG/TIG parser based on the DyALog logic programming environment [127] and on the Leff morphological and syntactic lexicon [118]. It relies on the notion of factorized grammar, themselves generated from a representation that lies at a higher level of abstraction, named Meta-Grammars [129]. At that level, linguistic generalizations can be expressed, which in turn makes it possible to transfer meta-grammars from one language to a closely related one. The hybrid TAG/TIG parser generator itself implements all kinds of parsing optimizations: lexicalization (in particular via hypertags), left-corner guiding, top/bottom feature analysis, TIG analysis (with multiple adjoining), and others.

Éric de La Clergerie has continued to improve the coverage, quality and efficiency of the French meta-grammar FRMG. On the EasyDev corpus (around 4000 sentences), parsing times have improved over 2011 from an average of 1.03s per sentence to 0.28s, coverage (in terms of sentences with full parses) has improved from 72.5% to 82.60%, and accuracy (in terms of f-measure over relations) from 64.54% to 68.28%.

A part of the accuracy gains comes from the addition of a new output format for FRMG, namely the CONLL format, allowing us to use the CONLL-based dependency version of the French Treebank (around 12K sentences) for training and evaluation. We also used new machine learning techniques to improve FRMG’s disambiguation algorithm, allowing us to combine heuristic based disambiguation rules (with manually provided weights) with more standard parsing features associated with automatically learned weights. More precisely, the idea was to study the efficiency of the disambiguation rules over the French treebank and to favor (resp. penalize) well-working (resp. bad working) rules by adjusting their weight, taking into account additional (and more standard) features. Using these techniques, on ftb6_3 test part, FRMG improved from a base accuracy of 82.31% (in terms of CONLL Labeled Attachment Score) to 84.54%. These gains resulting from a training over the French TreeBank have also been observed (with however a lesser impact) on the EasyDev corpus (using a different format and using a different evaluation metric).

6.2. Task-based evaluation of syntactic lexica: coupling FRMG with various resources

**Participants:** Éric Villemonte de La Clergerie, Benoît Sagot.

The FRMG symbolic parser was used for comparing the performances of various syntactic lexicons as sources of information for parsing. The idea is to convert syntactic lexica other than the Leff into the Leff’s format, i.e., turn them into Alexina lexicons, and then use the resulting lexica together with the FRMG grammar for producing several parsers. These parsers only differ by the lexical information they rely on. Preliminary results had already been obtained in 2009 [119], but were restricted to one external lexicon, namely Lexicon-Grammar tables, and only to verbal entries (other entries were gathered from the Leff when using Lexicon-Grammar-based verbal entries). However, conversion tools for other resources, such as Dicovalence [136], had already been developed, in the context of the development and improvement of the Leff. Moreover, the development of a new version of the Leff verbal entries.
Task-based evaluation results have been obtained on parsing with FRMG, showing that the Lefff performs better than both Lexique-Grammaire and DICOVALENCE (after conversion to the Alexina formalism) [48], [49]. The new version of the Lefff, mentioned above, leads for now to lower results than the current version, but its results are better than with Lexique-Grammaire or DICOVALENCE data, despite a significant increase of the average amount of entries per lemma. These results are satisfying both because they show that the Lefff is a useful resource for symbolic parsing, but also because they illustrate the relevance of converting other resources into the Alexina formalism, in order to merge the valuable linguistic information they contain — as done in the last years for improving the Lefff [85], [84], [86], [111], [87], [112].

6.3. Information extraction from corpora parsed with FRMG

Participants: Yayoi Nakamura-Delloye, Rosa Stern, Éric Villemonte de La Clergerie, Benoît Sagot.

Following previous experiments, in particular in the context of the FUI-funded project Scribo that ended in 2010, work has been achieved for extracting information from corpora parsed with FRMG.

In the context of the EDyLex project, we have proposed two pattern-based named entity extraction methods for ontology enrichment [36], [35]. The proposed methods are characterized by the use of entity relation patterns obtained by our unsupervised extraction method. These patterns correspond to syntactic paths that connect two named entities in dependency trees produced by FRMG. This work aims to take advantage of parsing benefits and also offers solutions for parsing disadvantage. The proposed methods are characterized by the use of entity relation patterns obtained by our unsupervised extraction method. These patterns correspond to syntactic paths that connect two named entities in dependency trees. This work aims to take advantage of parsing benefits and also offers solutions for parsing disadvantage.

We also developed a mechanism for integrating the results into a domain ontology, namely the ontology under deployment at the Agence France-Presse [37].

6.4. Advances in statistical parsing

Participants: Marie Candito, Benoît Crabbé, Djamé Seddah, Enrique Henestroza Anguiano.

6.4.1. Improving statistical dependency parsing

Alpage has provided state-of-the-art results for French statistical Parsing, adapting existing techniques for French, a richer morphological language than English, either for constituency parsing or dependency parsing. The Bonsai tool (see section 5.4) is available, that gathers preprocessing tools and models for dependency parsing French. We have innovated in the tuning of tagsets and the handling of unknown words. In the last years, Alpage has then contributed on four main points:

- conversion of the French Treebank [59] used as constituency training data into dependencies [72], the resulting treebank being used by several teams for dependency parsing;
- an original method to reduce lexical data sparseness and include coverage and robustness by replacing tokens by unsupervised word clusters or morphological clusters [69], [121], [73];
- all of our morphological clustering approaches were integrated into our parsing chains; data driven lemmatization required the adaptation of a state-of-the-art part-of-speech tagger and lemmatizer (Morfette [77]) based on a data-driven joint model benefiting the inclusion of external lexica such as the Lefff [121];
- a parser-agnostic postprocessing step, developed this year, which uses specialized models for dependency parse correction [30]; dependencies in an input parse tree are revised by selecting, for a given dependent, the best governor from within a small set of candidates, using a discriminative linear ranking model that includes a rich feature set that encodes syntactic structure in the input parse tree; the parse correction framework can correct attachments using either a generic model or specialized models tailored to difficult attachment types like coordination and pp-attachment; our

experiments have shown that parse correction, combining a generic model with specialized models for difficult attachment types, can successfully improve the quality of predicted parse trees output by several representative state-of-the-art dependency parsers for French.

- an adaptation of the above-mentioned technique of word clustering to the problem of adapting statistical parsers to different text domains. We show that in order to parse texts from a different domain than the one a statistical parser is trained on (namely to parse target domain text using a parser trained on indomain treebank), word clusters computed over a bridge corpus that couples indomain and target domain raw texts do improve parsing performance on target domain, without degrade performance on indomain texts (contrary to previous domain adaptation techniques). To evaluate these experiments, we use as target domain biomedical texts. We have supervised the manual syntactic annotation of a test corpus from the biomedical domain (European Public Assessment Reports concerning the marketing authorization of medicinal products).

Besides this line of work, it should be noted that two parsing models built around Stochastic Tree Insertion Grammars are currently under investigation: experiments have been conducted on Spinal TIGs. Moreover, we are still improving the TIG-based dependency parser MICA, developed in collaboration with University of Marseilles, Columbia university and AT&T (see section 5.5).

6.4.2. Functional labelling

Alpage worked towards the improvement of a functional labeller to be used as a post-parsing tool on an unfolded parse forest (as outputted e.g. by the Berkeley parser in the Bonsai architecture) using CRF models of various orders thereby extending the previous maximum entropy labeller designed in the team. The use of CRFs for modelling triggered a collaboration with Isabelle Tellier and JP Prost (LIFO, Orleans). The labeller implementation has been considerably improved and the accuracy of the labeller has improved as well on correct treebank trees. However we found out that the feature engineering work outweights the formal improvements since we were able to show that the use of higher order graphical models were not contributing significantly to improve an unstructured model. Our modest gains come mostly from feature engineering. Moreover we notice that combined with a constituent parser the labeller does not improve at all on constituent parsing output. The reason being that our current architecture for the Bonsai parser is sequential (which is unsatisfactory). Following experiments on n-best parsing outputs, we observe that the labeller can drastically improve on better parses where its input is indeed correct. This suggest investigating formulating constituent parsing and functional labelling as a joint task requiring to address serious efficiency issues. We intend to tackle the two drawbacks of our current architecture (sequential process, parse forest unfolding) by formulating constituent parsing as a joint task with functional labelling in the next few months.

6.4.3. Parsing spontaneous oral text

Alpage also got involved in parsing spontaneous oral text taken from ESTER 3 data (with overlaps) generated in the ANR ETAPE project in collaboration with A. Abeillé (LLF) with the aim of preannotating a seed for a future treebank of oral French which would considerably support work in experimental linguistics led in the Labex. He has also a collaboration set up with A. Abeillé, C. Gardent and C. Cerisara for ensuring interoperability accross ongoing efforts for producing oral treebanks for French. The way to carry out the task was by using a form of preprocessing of oral text to simulate a written entry to the Bonsai parser trained on written text. In the next few months we intend to test semi-supervised learning techniques to speed up the annotation process made by the LLF lab.

6.5. Named Entity Recognition and Entity Linking

Participants: Rosa Stern, Benoît Sagot.
Identifying named entities is a widely studied issue in Natural Language Processing, because named entities are crucial targets in information extraction or retrieval tasks, but also for preparing further NLP tasks (e.g., parsing). Therefore a vast amount of work has been published that is dedicated to named entity recognition, i.e., the task of identification of named entity mentions (spans of text denoting a named entity), and sometimes types. However, real-life applications need not only identify named entity mentions, but also know which real entity they refer to; this issue is addressed in tasks such as knowledge base population with entity resolution and linking, which require an inventory of entities is required prior to those tasks in order to constitute a reference.

6.5.1. Improvements of the Aleda entity database
Within the Alexina framework, we develop since 2012 the entity database Aleda [124], aimed at constituting such a reference. Aleda was first developed for French but is under development for English. Aleda is extracted automatically from Wikipedia and Geonames. It is used among others in the SxPipe processing chain and its NP named entity recognition, as well as in the NOMOS named entity linking system.

In 2011, major efforts have been made for improving the coverage, precision and richness of the French Aleda: improvements in the tool for creating an XML almost-raw-text version of the wikipedia, new method for identifying and typing entities among wikipedia articles, based on infoboxes and wikipedia categories, richer database structure for storing more detailed information about each entity, and many other improvements. A paper about these advances has been submitted to LREC 2012.

6.5.2. Cooperation of symbolic and statistical methods for named entity recognition and typing
Named entity recognition and typing is achieved both by symbolic and probabilistic systems. We have performed an experiment [24] for making the rule-based system NP, SxPipe’s high-precision named entity recognition system developed at Alpage on AFP news corpora and which relies on the Aleda named entity database, interact with LIANE, a high-recall probabilistic system developed by Frédéric Béchet and trained on oral transcriptions from the ESTER corpus. We have shown that a probabilistic system such as LIANE can be adapted to a new type of corpus in a non-supervised way thanks to large-scale corpora automatically annotated by NP. This adaptation does not require any additional manual annotation and illustrates the complementarity between numeric and symbolic techniques for tackling linguistic tasks.

6.5.3. Nomos, a statistical named entity linking system
For information extraction from news wires, entities such as persons, locations or organizations are especially relevant in a knowledge acquisition context. Through a process of named entity recognition and entity linking applied jointly, we aim at the extraction and complete identification of these relevant entities, which are meant to enrich textual content in the form of metadata. In order to store and access extracted knowledge in a structured and coherent way, we aim at populating an ontological reference base with these metadata. We have pursued our efforts in this direction, using an approach where NLP tools have early access to Linked Data resources and thus have the ability to produce metadata integrated in the Linked Data framework. In particular, we have studied how the entity linking process in this task must deal with noisy data, as opposed to the general case where only correct entity identification is provided.

We use the symbolic named entity recognition system NP, a component of SxPipe, and use it as a mention detection module. Its output is then processed through our entity linking system, which is based on a supervised model learnt from examples of linkings. Since our named entity recognition is not deterministic, as opposed to other entity linking tasks where the gold named entity recognition results are provided, it is configured to remain ambiguous and non-deterministic, i.e., its output preserves a number of ambiguities which are usually resolved at this level. In particular, no disambiguation is made in the cases of multiple possible mentions boundaries (e.g., \{Paris\}+(Hilton) vs. \{Paris Hilton\}). In order to cope with possible false mention matches, which should be discarded as linking queries, the named entity recognition output is made more ambiguous by adding a not-an-entity alternative to each mention detection. The entity linking module’s input therefore consists in multiple possible readings of sentences. For each reading, this module must perform entity linking on every possible entity mention by selecting their most probable matching entity. Competing readings are
then ranked according to the score of entities (or sequence of entities) ranked first in each of them. The reading with no entity should also receive a score in order to be included in the ranking. The motivation for this joint task lies in the frequent necessity of accessing contextual and referential information in order to complete an accurate named entity recognition; thus the part where named entity recognition usually resolves a number of ambiguities is left for the entity linking module, which uses contextual and referential information about entities.

We have realized a first implementation of our system, as well as experiments and evaluation results. In particular, when using knowledge about entities to perform entity linking, we discuss the usefulness of domain specific knowledge and the problem of domain adaptation.

6.6. Extending wordnets

Participants: Benoît Sagot, Marianna Apidianaki, Valérie Hanoka.

The WOLF (see section 5.9) is a freely available, automatically created wordnet for French, the biggest drawback of which has until now been the lack of general concepts that are typically expressed with highly polysemous vocabulary that is on the one hand the most valuable for applications in human language technologies but also the most difficult to add to wordnet accurately with automatic methods on the other. In collaboration with Darja Fišer (University of Ljubljana), we have developed a self-training-like technique for acquiring a classifier that is able to assign appropriate synset ids (i.e., senses) to new words, extracted from non-disambituated multilingual sources of lexical knowledge, such as Wiktionaries and Wikipedia [39], [40]. Automatic and manual evaluation shows high coverage as well as high quality of the resulting lexico-semantic repository. Another important advantage of the approach is that it is fully automatic and language-independent and can therefore be applied to any other language still lacking a wordnet. Indeed, it was applied to Slovene as well.

Other techniques were used as well and are the basis of various submitted conference papers. They rely, among others, on morphological derivation, on graph-based representation of highly multilingual lexicons extracted from numerous wiktionaries, and on automatically induced sense clusters.

6.7. Unsupervised lexical semantics

Participant: Marianna Apidianaki.

6.7.1. Unsupervised word sense induction and disambiguation

Word sense induction (WSI) is the task aimed at automatically identifying the senses of words in texts, without the need for handcrafted resources or annotated data. Up till now, most WSI algorithms extract the different senses of a word ‘locally’ on a per-word basis, i.e. the different senses for each word are determined separately. In collaboration with Tim van de Cruys, at Alpage in 2010, now at University of Cambridge [19], [50], we have compared the performance of such algorithms to a new algorithm that uses a ‘global’ approach, i.e. the different senses of a particular word are determined by comparing them to, and demarcating them from, the senses of other words in a full-blown word space model. The induction step and the disambiguation step are based on the same principle: words and contexts are mapped to a limited number of topical dimensions in a latent semantic word space. The intuition is that a particular sense is associated with a particular topic, so that different senses can be discriminated through their association with particular topical dimensions; in a similar vein, a particular instance of a word can be disambiguated by determining its most important topical dimensions. We evaluated our model on the SemEval-2010 word sense induction and disambiguation task. All systems that participated in this task use a local scheme for determining the different senses of a word. We obtain state-of-the-art results.
6.7.2. Unsupervised cross-lingual lexical substitution

Cross-Linguial Lexical Substitution (CLLS) is the task that aims at providing for a target word in context several alternative substitute words in another language. The proposed sets of translations may come from external resources or be extracted from textual data. In 2011, we have introduced a new approach for this task [18], namely the use of an unsupervised cross-lingual word-sense induction method. This method identifies the senses of words by clustering their translations according to their semantic similarity. We evaluated the impact of using clustering information for CLLS on the SemEval-2010 CLLS data set. Our system performs better on the ‘out-of-ten’ measure than the systems that participated in the SemEval task.


Participants: Pierre Magistry, Benoît Sagot.

For most languages using the Latin alphabet, tokenizing a text on spaces and punctuation marks is a good approximation of a segmentation into lexical units. Although this approximation hides many difficulties, they do not compare with those arising when dealing with languages that do not use spaces, such as Mandarin Chinese. Many segmentation systems have been proposed, some of them use linguistically motivated unsupervised algorithms. However, standard evaluation practices fail to account for some properties of such systems. New results [33] have shown that a simple model, based on an entropy-based reformulation of a language-independent hypothesis put forward by Harris in 1955, allows for segmenting a corpus and extracting a lexicon from the results. Tested on the Academia Sinica Corpus, our system allows for inducing a segmentation and a lexicon with good intrinsic properties and whose characteristics are similar to those of the lexicon underlying the manually-segmented corpus. Recent unpublished work using a slightly different model have improved these results. In parallel, preliminary experiments on other languages (Hindi, Singalese, Tamil, French) and original visualization techniques have already led to promising results.

6.9. Computational morphology

Participants: Benoît Sagot, Géraldine Walther.

Although computational morphology has been a topic of interest for Alpage for several years now, several new research topics have received attention in 2011, often in collaboration with morphologists from the Laboratoire de Linguistique Formelle (University Paris 7).

6.9.1. Inflectional morphology

Non-canonical inflection (suppletion, deponency, heteroclasis...) is extensively studied in theoretical approaches to morphology. However, these studies often lack practical implementations associated with large-scale lexica. Yet these are precisely the requirements for objective comparative studies on the complexity of morphological descriptions. We have shown [16], [43] how the Parsli model of inflectional morphology [132], which can represent many non-canonical phenomena, as well as a formalisation and an implementation thereof can be used to evaluate the complexity of competing morphological descriptions. After illustrating the properties of the model with data about French, Latin, Italian, Persian and Sorani Kurdish verbs and about noun classes from Croatian and Slovak we have conducted experiments on the complexity of four competing descriptions of French verbal inflection. The complexity is evaluated using the information-theoretic concept of description length. We show that the new concepts introduced in the model by the Parsli model enable reducing the complexity of morphological descriptions w.r.t. both traditional or more recent models.

6.9.2. Derivational morphology

This year, in relation with the ANR project EDyLex (see section 8.2.2), work has started targeted towards the acquisition of lexical information at the level of derivational morphology, both using semi- and non-supervised techniques.
Semi-supervised techniques have been used in a work dedicated to French denominal adjectives, for which we have implemented an automatic technique based on large-scale lexicons and corpora for extracting derivation links between base nouns and derived adjectives based on the same stem [46]. The resulting derivational lexicon, which is freely available, has already been partially manually validated. Future work include a full validation and adding denominal adjectives with a suppletive base. Unsupervised techniques have been used for extraction of derivational links that appear more systematically, although their definition is less linguistically motivated as such [51].

6.9.3. Morphological issues concerning loan words

Also in the context of the ANR project EDyLex (see section 8.2.2), we have carried out a preliminary study on the morphological issues raised by borrowing phenomena, concerning in particular French nouns and verbs borrowed from English [52]. Using techniques that are similar to those used on derivational morphology, we have extracted a significant amount of loan words from a large raw corpus. We have proposed a model of the borrowing phenomenon, that takes into account graphemic (spelling), phonetic and morphological variability.

6.10. Allophony and word segmentation in language acquisition models

Participants: Luc Boruta, Benoît Crabbé.

Allophonic rules are responsible for the great variety in phoneme realizations. Infants cannot reliably infer abstract word representations without knowledge of their native allophonic grammar. We have explored the hypothesis that some properties of infants’ input, referred to as indicators, are correlated with allophony. First, we provide an extensive evaluation of individual indicators that rely on distributional or lexical information. This evaluation relies on a phonetically transcribed corpus, generated automatically from a phonemically transcribed English, French and Japanese child-directed corpus. As such corpora do not exist as such, we used automatically extracted allophonic grammars of various sizes leading to various granularity levels, using our own allophonic rule extraction algorithm [57]. Then, we present a first evaluation of the combination of indicators of different types, considering both logical and numerical combinations schemes [23]. Though distributional and lexical indicators are not redundant, straightforward combinations do not outperform individual indicators.

Models of the acquisition of word segmentation are typically evaluated using phonemically transcribed corpora. Accordingly, they implicitly assume that children know how to undo phonetic variation when they learn to extract words from speech. Moreover, whereas models of language acquisition should perform similarly across languages, evaluation is often limited to English samples. Using the phonetically annotated corpora described above, that cover three typologically different languages, we evaluated the performance of state-of-the-art statistical models given inputs where phonetic variation has not been reduced. We have measured segmentation robustness across different levels of segmental variation, simulating systematic allophonic variation or errors in phoneme recognition. We have shown that these models do not resist an increase in such variations and do not generalize to typologically different languages. From the perspective of early language acquisition, the results strengthen the hypothesis according to which phonological knowledge is acquired in large part before the construction of a lexicon.

6.11. Modelling the acquisition of syntactic categories by children

Participant: Benoît Crabbé.

B. Crabbé co-supervised A. Gutman for an M2 thesis (MPRI) in collaboration with A. Christophe (LSCP/ENS) in the domain of psycholinguistic modelling. The topic was concerned with modelling and implementing psycholinguistically motivated models of language treatment and acquisition. Contrary to classical Natural Language Processing applications, the main aim was not to create engineering solutions to language related tasks, but rather to test and develop psycholinguistic theories. In this context, the study was concerned with the question of learning word categories, such as the categories of Noun and Verb. It is established experimentally that 2-year-old children can identify novel nouns and verbs. It has been suggested that this can be done
using distributional cues as well as prosodic cues. While the plain distributional hypothesis had been tested quite extensively, the importance of prosodic cues had not been addressed in a computational simulation. We provided a formulation for modelling this hypothesis using unsupervised and semi-supervised forms of bayesian learning (EM) both offline and online.

6.12. Modelling and extracting discourse structures

Participants: Laurence Danlos, Charlotte Roze.

6.12.1. Cross-lingual lexical semantics of discourse connectives

Discourse connectives are words or phrases that indicate senses holding between two spans of text. The theoretical approaches accounting for these senses, such as text coherence, cohesion, or rhetorical structure theory, share at least one common feature: they acknowledge that many connectives can indicate different senses depending on their context. Depending on its sense, the translation of a connective into another language can vary greatly, either using an equivalent connective, or using a different construction or even no explicit connective at all.

On the basis of data provided by the bilingual concordancer TransSearch which propose statistical word alignment [64], [53] made a semi-manual annotation of the English translation of two French connectives ("en effet" and "alors que"). The results of this annotation show that the translations of these connectives do not correspond to the “transpots” identified by TransSearch and even less to the translations proposed in bilingual dictionaries.

The conclusions of this work were presented at an European workshop organized by the project COMTIS\(^5\), and some members decide to use our technic for other connectives and other aligned corpora (e.g. Europarl).

6.12.2. Discourse relations inference rules

In 2011 we have developed a new methodology for building discourse relations inference rules, to be integrated into an algebra of these relations [54], [38]. The construction of such an algebra has as main objective the improvement of the comparison of discourse structures within the evaluation of discourse annotations and the creation of a gold-standard corpus. The inference rules can also help detecting inconsistencies in discourse structures, in order to improve human or machine annotation. The premises of rules already studied lead to the formulation of inference rules, established by the theoretical definition of discourse relations, manually constructed data and extracted data. By manually annotating discourses, we also compute inference probabilities. We have illustrated the adopted methodology taking as theoretical background the Segmented Discourse Representation Theory [60].

6.12.3. Discourse structure and factivity

Discursive annotations proposed in theories of discourse such as RST (Rhetorical Structure Theory) or SDRT (Segmented Representation Theory Discourse) have the advantage of building a global discourse structure linking all the information in a text. Discursive annotations proposed in PDTB (Penn Discourse Tree Bank) have the advantage of identifying the "source" of each information – thereby answering to questions such as who says or thinks what?

In collaboration with Owen Rambow (Columbia University), we have proposed [26], [28] a unified approach for discursive annotations combining the strengths of these two streams of research. This unified approach relies crucially on factivity information, as encoded in the English corpus FactBank. We intend to pursue this avenue of research by initiating in 2012 the development of a French FactBank.

6.13. Modelling and extracting temporal structures

Participants: Pascal Denis, Philippe Muller.

\(^5\) http://www.idiap.ch/project/comtis
Temporal information has been the focus of recent attention in information extraction. An elegant approach to learning temporal orderings from texts is to formulate this problem as a constraint optimization problem, which can be then given an exact solution using Integer Linear Programming. This works well for cases where the number of possible relations between temporal entities is restricted to the mere precedence relation, but becomes impractical when considering all possible interval relations.

We have proposed this year two innovations [29], inspired from work on temporal reasoning, that control this combinatorial blow-up, therefore rendering an exact ILP inference viable in the general case. First, we propose to translate the network of constraints from temporal intervals to their end-points, to handle a drastically smaller set of constraints, while preserving the same temporal information. Second, we have show that additional efficiency is gained by enforcing coherence on particular subsets of the entire temporal graphs. We evaluate these innovations through various experiments on TimeBank 1.2 using standard evaluation metrics, and compare our ILP formulations with various baselines and oracle systems.

The evaluation of temporal information extraction, i.e., the comparison of two annotations of a given text, is also a scientific challenge. This is because relations between events in a story are intrinsically interdependent and cannot be evaluated separately. A proper evaluation measure is also crucial in the context of a machine learning approach to the problem. Finding a common comparison referent at the text level is not obvious, and we have argued, in collaboration with Xavier Tannier (LIMSI), in favor of a shift from event based measures to measures on a unique textual object, a minimal underlying temporal graph, or more formally the transitive reduction of the graph of relations between event boundaries [15].


Participants: Emmanuel Lassalle, Pascal Denis.

Bridging descriptions are a special kind of anaphora whose interpretation requires not only identifying an antecedent, but also inferring a specific relation linking it to the anaphor. The resolution of bridging anaphora represents a very challenging task in discourse processing. It is considerably much harder than standard coreferential anaphora resolution for which shallow predictors (like distance, string matching, or morphosyntactic agreement) have been shown to be rather effective. Part of the challenge is due to an important information bottleneck. Lexical resources like WordNet are still too poor and uneven in coverage to provide a realistic solution. In turn, more recent approaches to bridging resolution have turned to web-based extraction methods. To date, the most complete and best-performing approach combines focus and lexical distance predictors using machine learning techniques [105].

We have focused on mereological bridging anaphora (that is, cases wherein the inferred relation is a part-whole relation). Moreover, we have worked on French, a language for which current lexical resources have a very low coverage. The system, presented in [32] is similar to a system developed for English [105], but it was enriched to integrate meronymic information extracted automatically from both web queries and raw text using syntactic patterns. Through various experiments on the DEDE corpus [78], we show that although still mediocre the performance of our system compare favorably to those obtained for English by the above-mentioned system. In addition, our evaluation indicates that the different meronym extraction methods have a cumulative effect, but that the text pattern-based extraction method is more robust and leads to higher accuracy than the Web-based approach.

6.15. Statistical models of word order in French

Participants: Juliette Thuilier, Benoît Crabbé.

We study the problem of choice in the ordering of French words using statistical models along the lines of [66] and [67]. This work aims at describing and model preferences in syntax, bringing additional elements to Bresnan’s thesis, according to which the syntactic competence of human beings can be largely simulated by probabilistic models. We previously investigated the relative position of attributive adjectives with respect to the noun.

---

6An illustrative English example is is the following discourse: *The car will not move. The engine is broken.*
This year, we mainly studied the problem of the relative ordering of postverbal complements. The focus of this investigation is the relative order of direct object and indirect object of French ditransitive verbs. The first part of this work is based on corpora data that we extracted from two journalistic corpora (French Tree Bank and Est-Républicain) and a radio corpus (ESTER). These data were manually annotated and validated for semantic categories (animacy and semantic class of the ditransitive verb). Based on these data, we built statistical models showing that the relative length of complements and verbal lemmas are the most important factors, and that, differently from English or German, categories as animacy or definiteness seem to play no role in the relative ordering.

In collaboration with Anne Abeillé (Laboratoire de Linguistique Formelle, Université Paris 7), we extended our corpora study with psycholinguistic questionnaires, in order to show that statistical models are reflecting some linguistic knowledge of French speakers. The preliminary results confirm that animacy is not a relevant factor in ordering French complements.

As regards to corpus work, we are extending the database with spontaneous speech corpora (CORAL-ROM and CORPAIX) and a wider variety of verbal lemmas, in order to enhance sample representativeness and statistical modelling. In a crosslinguistic perspective, we plan to strengthen the comparison with the constraints observed in other languages such as English or German.

As can be seen from the outline above, this line of research brings us closer to cognitive sciences. We hope in the very long run that these investigations will bring new insights on the design of probabilistic parsers or generators. In NLP the framework that is closest to implementing construction grammar is Data Oriented Parsing.

### 6.16. Assessing the Amazon Mechanical Turk plateform

**Participant:** Benoît Sagot.

In collaboration with Gilles Adda and Joseph Mariani from LIMSI and with Karên Fort from INIST, we have assessed some crowdsourced microworking systems and especially Amazon Mechanical Turk, the use of which has been steadily growing in language processing in the past few years [41], [17]. According to the mainstream opinion expressed in the articles of the domain, this type of on-line working platforms allows to develop very quickly all sorts of quality language resources, for a very low price, by people doing that as a hobby or wanting some extra cash. We have demonstrated that the situation is far from being that ideal, be it from the point of view of quality, price, workers’ status or ethics and bring back to mind already existing or proposed alternatives. Our goal was threefold:

- to inform researchers, so that they can make their own choices with all the elements of the reflection in mind,
- to ask for help from funding agencies and scientific associations, and develop alternatives,
- to propose practical and organizational solutions in order to improve new language resources development, while limiting the risks of ethical and legal issues without letting go price or quality.

### 6.17. Finite state formalisms for Egyptian Hieroglyphic transliteration

**Participant:** François Barthélémy.

The task of transliterating an Egyptian Hieroglyphic text into the latin alphabet was studied [20], as a model problem to compare two finite-state formalisms: the first one is a cascade of binary transducers; the second one is a class of multitape transducers expressing simultaneous constraints, implemented using the Karamel language [62]. The two systems were compared regarding their expressivity and readability.

The first system tends to produce smaller machines, but is more tricky. On the other hand, the Karamel language provides a more abstract description of the forms, using an explicit tree structure and separating the different pieces of information on different tapes, according to semantic criteria. But the Karamel machine is much larger. Karamel is a high-level declarative formalism whereas non contextual rewrite rules are an efficient low-level language.
5. New Results

5.1. Introduction

This year we obtained new results in our three sub-objectives and also related to Focuslab platform and software valorization:

1. **Sub-objective 1 - Mining for Knowledge Discovery in Information Systems:** this we get five results (with one achieved Ph.D thesis).

   Let us note that six past works on this sub-objective described in previous AxIS annual reports have been published this year as articles in international journals ([22],[11]) or conferences, one in a national journal [50], two in a french-speaking conference [35],[46], one book [20] and one book chapter [52] at international level. Indeed
   - The work in 2009 on mining data streams by Marascu in her thesis [96] has been published in [11] with more details in the algorithms and in the experiments.
   - The work published in 2008 on discovering frequent behaviors [107] has been published in [22] with more details in the algorithms and in the experiments.
   - Our previous work on satellite image mining in 2010 [99] has been published in French at EGC [35].
   - Our past work on Functional data analysis involving data described by regular functions rather than by a finite number of real valued variables has been published as a scientific book chapter [52]. In this paper we propose to use a clustering approach that targets variables rather than individual to design a piecewise constant representation of a set of functions. The contiguity constraint induced by the functional nature of the variables allows a polynomial complexity algorithm to give the optimal solution.
   - In the context of the WRUM project (Morocco) and Zemmouri’s PhD thesis, we have a long paper accepted at JFO 2011 related to past works (2010) on how to integrate domain knowledge in a multi-view KDD process [46].

2. **Sub-objective 2 - Information and Social Networks Mining for Supporting Information Retrieval:** Three results (with one achieved Ph-D thesis). Let us remind the best paper [15] obtained By E. Smirnova at ECIR 2011 for her research on expert finding.

3. **Sub-objective 3 - Multidisciplinary Research For Supporting User Oriented Innovation:** this interdisciplinary research is dedicated to the design, tailoring and refinement of methodologies and tools for a better users’ involvement in innovation processes. We have seven results this year.

Concerning our activity in terms of **FocusLab Experimental Platform and Software**, a) we first applied ATWUEDA on another context of evolutive data (on system monitoring data at EDF) which is different of Web usage data) to show the genericity of the approach [9], and b) we develop a Web-based version of the FocusLab experimental platform for analysis usage data (hardware and software parts).

5.2. Mining for Knowledge Discovery in Information Systems

5.2.1. Mining Data Streams: Clustering and Pattern extraction

**Participant:** Chongsheng Zhang.
In Zhang’s thesis [19] (supervised by F. Masseglia), which was partially founded by ANR MIDAS (cf. 6.2.1), we present our study of the management and mining issues on data streams with evolving tuples, caused by model updates or tuple revisions. For instance, in an online auction system where bids on auction items are streaming, it is possible that some users may bid for more than one item within the user-specified time interval. As a result, the profiles of the users can be updated or revised in such applications. Data streams having evolving tuples bring new challenges as well as research opportunity. In this work, he develops novel and efficient models and methods for managing and mining data streams with evolving tuples. (I) To model data streams with evolving tuples, we propose the Anti-Bouncing Streaming model (ABS) for usage streams. ABS fits data streams with evolving tuples and it enables methods for processing of data streams to handle tuple updates or revisions. (II) To find frequent itemsets from data streams with evolving tuples over pane-based sliding windows, we conduct theoretical analysis and propose theorems which can avoid scanning the past slides to check for possible itemsets that may become frequent. We also design novel data structures which can manage the data streams with evolving tuples efficiently and facilitate the mining of frequent itemsets. Moreover, we devise an efficient counting algorithm to verify the frequentness of the candidate frequent itemsets. We also propose two running frameworks for this problem. (III) To extract important feature set from data streams (including the ones with evolving tuples), based upon ABS, we devise the streaming feature set selection algorithm for data streams which is the first in the literature. This method is based on information theory to extract the informative feature sets. To further accelerate the extraction of the most informative feature set from high-dimensional data, we propose a framework that reduces the huge search space to a rather small subset while still guarantee the quality of the discovered feature sets.

In 2011, Chongsheng Zhang has mainly worked on a data stream mining method, intending to extract frequent itemsets. This method has not been published yet and is described in Chapter 5 (page 79) of his thesis document [19].

5.2.2. Clustering on Multiple Dissimilarity Matrices

Participants: Yves Lechevallier, Francisco de A.T. de Carvalho, Thierry Despeyroux, Alessandra Silva Anyzewski.

In [23] we introduce hard clustering algorithms that are able to partitioning objects taking into account simultaneously their relational descriptions given by multiple dissimilarity matrices [49]. The aim is to obtain a collaborative role of the different dissimilarity matrices in order to obtain a final consensus partition. These matrices could have been generated using different sets of variables and a fixed dissimilarity function or using a fixed set of variables and different dissimilarity functions, or using different sets of variables and dissimilarity functions.

These methods, which are based on the dynamic hard clustering algorithm for relational data as well as on the dynamic clustering algorithm based on adaptive distances, are designed to furnish a partition and a prototype for each cluster as well as to learn a relevance weight for each dissimilarity matrix by optimizing an adequacy criterion that measures the fitting between clusters and their representatives.

These relevance weights change at each algorithm iteration and can either be the same for all clusters or different from one cluster to another. The usefulness of these partitioning hard clustering algorithms are shown on two time trajectory real world datasets.

5.2.3. Clustering of Constrained Symbolic Data

Participants: Marc Csernel, Francisco de A.T. de Carvalho.

In the context of our FACEPE collaboration with Brazil (cf. section 6.4.3.1), we have presented a method which allows clustering of symbolic descriptions constrained by presence rules in a polynomial time instead of a combinatorial one. This method allows to deal with “false missing values”. Such a method can be applied on various classification problems [26].

5.2.4. Web Page Clustering based on a Community Detection Algorithm

Participants: Yves Lechevallier, Yacine Slimani.
Extracting knowledge from Web user’s access data in Web Usage Mining (WUM) process is a challenging task that is continuing to gain importance as the size of the web and its user-base increase. That is why meaningful methods have been proposed in the literature in order to understand the behaviour of the user in the web and improve the access modes to information. In this work [42], we are interested in the analysis of the user browsing behavior. The objective is to understand the navigational practices of users (teachers, students and administrative staff). First we clean the data by removing irrelevant information and noise. During the second step, remaining data are arranged in a coherent way in order to identify user sessions. After we defined a new approach [42] of knowledge extraction. This approach treats the data resulting from the preprocessing phase (first and second steps) as being a set of communities. Our approach extends the Modularity measure, proposed by Newman and Girvan [97], in the Web Mining context in order to benefit from their classifying capacity in the communities discovery.

This work is done in collaboration with the LRIA laboratory – Université Ferhat Abbas, Sétif, Algérie

5.2.5. Critical Edition of Sanskrit Texts

Participants: Marc Csernel, Nicolas Béchet, Ehab Hassan, Yves Lechevallier.

New progresses concerning the computer assisted elaboration of Sanskrit texts have been made. First Nicolas Béchet and Marc Csernel have worked on the problem of moved texts. After an alignment between two versions of the texts, we discover that some parts of the text apppears to have been moved according to the technics developed in [48]. Until now, we were not able to discover when a text has been moved in a manuscript.

Now using a words-grams technique proposed in [48], we were able to obtain quite good results on the moved texts problem and we were able to optimize the different possible parameters. A paper on the subject has been submitted to the Cicling 2012 conference (http://www.cicling.org/2012/).

After the new treatment related to the moved text problem, we need to provide an interactive display of the critical edition. During his internship, Ehab Hassan has been working on the subject and obtained good results. These results need to be deeply examined by Sanskritists to see if they always fulfill their needs.

5.3. Information and Social Networks Mining for Supporting Information Retrieval

5.3.1. Clustering of Relational Data and Social Network Data

Participants: Yves Lechevallier, Amine Louati.

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

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

In the enterprise context, a considerable amount of information is stored in relational databases. Therefore, relational database can be a rich source to extract social network. The extracted network has in general a huge size which makes its analyses and visualization difficult tasks. In [45], we propose a social network extraction approach from relational database.

Often, the network has a large size which makes its analysis and visualization difficult. The aggregation step is a necessary task, so we offer [33] and [32] an aggregation step based on the k-SNAP algorithm [109] that produces a summary graph by grouping nodes based on attributes and relationships selected by the user.

This work is done in collaboration with Marie-Aude Aufaure, head of the Business Intelligence Team, Ecole Centrale Paris, MAS Laboratory.
5.3.2. Networks Solutions for Expert Finding and People Name Disambiguation

**Participants:** Elena Smirnova, Yi-Ling Kuo, Brigitte Trousse.

The task of finding people who are experts on a given topic has cently attracted close attention. State-of-the-art expert finding algorithms uncover knowledge areas of candidate experts based on textual content of associated documents. While powerful, these models ignore social structure that might be available. Therefore, we develop a Bayesian hierarchical model for expert finding that accounts for both content and social relationships. The model assumes that social links are determined by expertise similarity between candidates. The results of EGC experiments on UvT expert collection have demonstrated the effectiveness of our algorithm [43].

E. Smirnova visited Intellius, people search technology company (Aug 8 - Oct 5, 2011): the goal of this visit was to validate the research on expert finding in social networks on real dataset and further advance it. As a real dataset, we have taken a sample of United States LinkedIn public profiles. We built an organizational network by connecting a LinkedIn user and his colleagues at different workplaces. We also constructed a geographical network from user’s current location in the United States. We used Amazon’s Mechanical Turk framework (http://aws.amazon.com/code/923) to collect user-oriented judgements for model evaluation. We found that the user-oriented model is statistically significantly preferred to the baseline model on 72.5% of queries.

Her work on name disambiguation done in 2010 has been integrated in an article related to the problem of quick detection of top-k Personalized PageRank (PPR) in [24]. The effectiveness of the chosen approach based on Monte Carlo methods for quick detection of top-k PPR lists has been demonstrated on the Web and Wikipedia graphs.

Yi-Ling Kuo during her internship has worked on Person Name Disambiguation and started by managing the analysis of the very huge Yahoo! Web graph.

This topic has been done in the context of Smirnova’s thesis [18] which has been defended on december 15 (thesis supervised by B. Trousse (AxIS) and K.Avrachenkov (Maestro)).

5.3.3. Towards an On-Line Analysis of Tweets Processing

**Participant:** Nicolas Béchet.

Tweets exchanged over the Internet represent an important source of information, even if their characteristics make them difficult to analyze (a maximum of 140 characters, etc.). In [25], we define a data warehouse model to analyze large volumes of tweets by proposing measures relevant in the context of knowledge discovery. The use of data warehouses as a tool for the storage and analysis of textual documents is not new but current measures are not well-suited to the specificities of the manipulated data. We also propose a new way for extracting the context of a concept in a hierarchy. Experiments carried out on real data underline the relevance of our proposal.

This work is done inside a collaboration with LIRMM and CEMAGREF.

5.4. Multidisciplinary Research For Supporting User Oriented Innovation

5.4.1. Usability Design and Evaluation Methods

**Participants:** Dominique Scapin, Yves Lechevallier, Pascal Marie-Dessoude, Claudia Detraux.

We pursued our work on articulation of usage mining approach and human factors expertise for the design and evaluation of information systems. Namely, collaborative clustering techniques were used to analyze data issued from users via a card sorting technique, with respect to an a priori (“expert”) clustering. Considering the difficulties that people have in managing large information sets in their everyday life, for either professional or non-professional purposes (administration, social relationships, leisure, etc.), our recent research focuses on personal information space for which information bits are currently scattered many places.
In this PIMs field, there is little research with a user-centric approach, with the view that users-based knowledge might help specifying computer-based tools and a state-of-the-art [37] showed little work specifically on usability. Studies address a variety of questions from user needs to accessibility (including studies on older people) or user acceptance, among others. In the context of user-centered and long-term studies to understand the evolution of user information practice, we looked in a study at the intuitive way people organize their personal information, with or without computer systems, in order to help the design of future systems. Also, we recently surveyed 15 tools that claim to support personal information management.

5.4.2. Living Lab Landscape

Participants: Marc Pallot, Brigitte Trousse, Bernard Senach, Dominique Scapin.

In order to provide to the research community a comprehensive landscape of research streams in the Living Lab domain, we launched a study on the state-of-the-art about the ubiquitous notion of User Experience. During this continuous study, a landscape [34], [69] has progressively emerged that we organized through 4 main axes: focus granularity (individual/group), user’s role in the design process (observed subject/value creator), collaboration style (structured/unstructured), and evaluation purpose (reliability/adoptability). Our landscape of research streams has been used by Finnish colleagues who conducted an empirical study on the use of the Living Lab research domain landscape as a tool for assessing the maturity level of 16 Finnish Living Labs [36]. The Living Lab research domain Landscape has allowed the study team to identify four categories of Living Labs.

5.4.3. Future Internet Domain Landscape

Participants: Marc Pallot, Brigitte Trousse, Bernard Senach.

There are many different Internet research areas and corresponding technologies that were already investigated, experimented and progressively deployed such as peer-to-peer, autonomous, cognitive and ad hoc networking, that have already demonstrated how to improve network performance and user experience. Peer-to-peer networking for large-scale distributed systems and widely used applications has proved both the feasibility and economic potential for delivering services to millions of users. Others emerged more recently in the context of the future Internet (FI), such as Cloud Computing for transparently sharing among users scalable elastic resources over a limitless network. As it remains difficult to visualise the conceptual evolution and articulate the various Internet research areas, we conducted a study for identifying the appropriate concepts that could populate the FI domain landscape [69], [71] over three different periods of time (1990-1999, 2000-2005 and 2006-2011). Several INRIA research teams are involved in FIRE (Future Internet Research Experimentation) Testbed projects, namely: PlanetLab, OneLab, TEFIS, SensLAB, and BonFIRE whose scientific leaders were interviewed during the development of the FI domain landscape. Four dimensions were used for landscaping the Future Internet research domain: evolution approaches (from incremental evolution design to Clean Slate re-design or radical evolution), Internet routing (from the basic data packet delivery towards more sophisticated content distribution and retrieval capacities such as content Centric Networking), network type (from wired communication to wireless communication networks), evolution trend (from computer network towards network computing). The resulting tentative landscape of FI research areas shown in Figure 1 is intended to provide a faster and broader understanding of the different Internet research streams and related topics.

5.4.4. Future of Internet and User-open innovation for Smart Cities

Participants: Caroline Tiffon, Marc Pallot, Brigitte Trousse, Bernard Senach.

The goal of the Fireball project is to bring together three different constituencies: user driven open innovation, Future Internet, and Smart Cities [34], [39]. It aims at defining a roadmap [29], based on analysis of needs, opportunities and gaps, to benefit a wide scale implementation of the methodologies and concepts elaborated. A first objective in the project was to get a clear picture of the state-of-the-art in each domain. During the review, progressively emerges a landscape [69] that we organized along 4 main axes: wiring (wired/wireless), user’s role (subject/actor), Internet evolution approach (structured/unstructured), evaluation purpose (reliability/adoptability). A large variety of FI research have been engaged. If initial efforts in Future
Internet research have been directed towards the goal of providing the technical infrastructure supporting the next network generation, a rising trend in this research field is to consider now a higher level layer, the layer of services.

5.4.5. Method and Tool for Selection of Open Innovation Software Tools

Participants: Mylène Leitzelman, Brigitte Trousse.

In spite of an important number of tools supporting open innovation, there is few comparative evaluation and no grid or evaluation criteria helping to choose a product. A 2011 review in the Computer Aided Innovation field provides a large overview of available tools in relation with a wide range of innovation cycle features.

This useful top-down categorization approach is of little help to choose a specific tool. To find the best OI tool supporting idea exchanges among a community of participants, we built an exploratory method on the Web and we elaborate an assessment grid of OI tools based on the QSOS method which is a method designed to qualify, select and compare free and open source software in an objective, traceable and argued way. It publicly available under the terms of the GNU Free Documentation License. In our QSOS method, evaluation criteria are organized in a tree-hierarchy grid with, a scoring method procedure of each tree-leaf criteria (from 0 - not covered to 2 - completely covered). To achieve the construction of the OI tools criteria assessment grid (first step), we used different mining tools for Web crawling, network analysis, criteria classification and from the 29 top rated OI tools, we finally selected 6 of them from which we extracted the tree-map categorization used to build the reference software criteria sheet. In the following steps, after appropriate weighting, we used the provided OS3 Web application and we were able to compare 4 top selected OI tools.

To support our method, we developed a QSOS-based OI Grid for supporting the OI tool selection. The QSOS Grid to compare OI tools, which is in an XML format, has been translated into a MindMap. As a first mock-up, the QSOS Web Interface O3S will be installed at the beginning of 2012 on our server. For the future, we...
will extend our own QSOS comparative method to other softwares and to other categories interesting for the Living Lab community.

![Figure 2. Radar Positioning of Selected Open Innovation Tools](image)

### 5.4.6. Extension of Methods for Idea Generation Process

**Participants:** Anne-Laure Negri, Caroline Tiffon, Bernard Senach, Brigitte Trousse.

Internet of Things (IoT) is supposed to be a physical world where everyday objects, rooms and machines are connected to one another and to the larger digital world. In this web of people and objects, individuals as well as things will have their own unique URL and, according to interactions between all these entities, social networks will be articulated with a ring of connected objects. This mash up of "hyper groups" and "hyper objects" will be the next step towards a deeper level of automation in which the user interface has vanished, the explicit control over the world is no longer the rule and where the users will have to understand the dynamic changes of their environment in relation to their social interactions and to their physical behaviour. The design of IoT based services therefore raises many challenges related to the user experience and requires a deep understanding of users’ needs in their real life environment as well as many field experiments; this is why the Living Lab approach appears as the qualification device which has been lacking so far. As described in a paper accepted for ServDes2012 conference (February 2012) among 80 papers submitted, we developed and tested two methods designed for Internet of Things service idea generation: GenIOT (Generative Technique for the Internet of Things) and Aloha! (Animation Lens: Object/Human Actoring!) methods.

- **The GenIOT method** aims at providing citizen a tool for identifying and sharing examples of situations where they may benefit from an Internet of Things based service. Indeed the objective is to help the introspective effort of analysing one’s own daily behavior and unveil situated data in order to develop grounded internet of things services ideation. Fake sensors are placed by the participants in their daily environment and pictures of the sensors in situ are shared and tagged on a collaborative platform. Other participants are invited to discuss online the ideas behind the pictures. Gamification rules are implemented in the platform in order to enhance collaboration and participation.

- **The Aloha! method** tackles another creative dimension as it asks participants to role or play characters or objects and bodystorm a collaborative scenario meeting the characters need. Participants report it to be a fun and effective method, alternative to traditional brainstorming and integrating serendipity as a creative asset.

For the exploration step in the context of ELLIOT, we developed NiceAir, an android mock-up for visualizing air quality data on Google map in Nice as well as some interest points (free bikes, free park places, bus stop, etc.). Such a mock-up will be available in the co-creation step of green services in the ELLIOT project and during the European Mobility Week.
5.4.7. **Mock-ups for two innovation processes (exploration step)**

**Participants:** Lucile Gramusset, Guillaume Pilot, Mohamed Gaieb, Bernard Senach, Brigitte Trousse.

In the context of two contracts (TICTAC and ELLIOT) related to user oriented innovation process, we have developed two mock-ups:

- Based on the user feedback from the first experiment in TICTAC (cf. section 6.1.1), we decided to develop a mock-up **MobilTIC** of a real-time information service related to public transportation (Envibus & SNCF) usable for any smartphone with internet access for Sophia-Antipolis workers. We designed a Web interface the most simple based on PHP/Mysql technologies and accessible from a computer or 3G mobile phone. Usage analysis of MobilTIC has been anticipated by providing rich logs. An exploration task with citizen is planned in January 2012 with a new experiment with an improved version.

- A first mock-up called **Nice Air** has been developed in the context of ELLIOT (cf. section 6.3.1.1) by L. Gramusset and M. Gaieb on android smartphone for providing information related to air quality and noise on a map of the area of Nice Cote d’Azur with some interest points such as bus stops and Vélib parkings.

5.5. **FocusLab Experimental Platform and Software**

5.5.1. **ATWUEDA based Clustering Approach for System Monitoring**

**Participant:** Yves Lechevallier.

Progressive advances in hardware and software technologies have enabled the production and storage of system monitoring data streams in a wide range of fields (e.g. telecommunications, sensor networks, etc.). Traditional clustering methods are unable to deal with data of such a voluminous and dynamic nature. In this work [51], we propose an efficient clustering approach (ATWUEDA) for monitoring massive time-changing data streams. This work considers a real case study on condition monitoring data streams of an electric power plant provided by EDF.

This work is done in collaboration with Alzennyr Da Silva of BILab laboratory (Telecom ParisTech and EDF R&D Common Laboratory)

5.5.2. **FocusLab Experimental Platform (CPER Telius 2008-2012)**

**Participants:** Xavier Augros, Mohamed Gaieb, Brigitte Trousse, Yves Lechevallier.

The FocusLab platform aims to be a major delivery mechanism of previous and current work in AxIS. It is a way to make methodological contributions (including software) available for the scientific community, but also a way for stimulating further research. This work has slowly started on the software part due to the absence of human resource funding and due to the absence of engineers in the team until 2010. Mid 2011, we started the specification and the development of a first version of a platform with three parts (hardware, software and methods) with the arrival of two engineers on the ELLIOT and TIC TAC contracts. A first version is available since the end of september as a Web portal and a second one is planned for the end of 2012 with advanced features for the software part (cf. service oriented platform, SOA architecture and interoperability).

Related to the software part we are in the process of developing several AxIS methods as Web services: we started with SCDS (cf. 4.2.3) which was demonstrated in the context of the MIDAS project on two applications (Orange labs mobile portal and vehicle trajectories) and in the ELLIOT platform (linked to the San rafaelle Hospital media use case) at the first review meeting (cf. 6.3.1.1) applied on data issued from San Rafaello Hospital use case. ATWUEDA (cf. 4.3.2), GEAR (Marascu’s thesis) are under development as web services.

Our work on mining evolutive data (ATWUEDA) and data streams (such as SCDS) have been used in real applications in the context of Internet of things and sensors: ATWUEDA (system monitoring for EDF (cf. section 5.5.1) and SDMS (cf. section 4.2.3).
6. New Results

6.1. High speed autonomous driving on roads

Participants: Fawzi Nashashibi, Laurent Bouraoui, Paulo Lopes Resende.

In this exceptional year, IMARA-INRIA had a very busy agenda with several scheduled demonstrations in the frameworks of several European projects, especially HAVEit, CityMobil and CityNetMobil. The HAVEit project final event took place at Boras (Sweden) in June. IMARA was tightly involved in the development of the Joint Demonstrator called FASCar-II. In this demonstrator, IMARA was responsible for the development of the Co-Pilot system. This is the main decisional system that handles the manoeuvres to be executed taking into account the multisensor data fusion sub-system, the driver monitoring system and the infrastructure (data provided by the infrastructure-to-vehicle telecommunications). The system also generates the trajectories to be executed as well the control-command laws to be send to the actuators \[ 40 \] and provides passive or active assistance to the driver according to the active automation level.

![Image of Co-Pilot system in Joint system architecture and HAVEit’s FASCar demonstrator](image)

Figure 1. The Co-Pilot system in the Joint system architecture and HAVEit’s FASCar demonstrator

In parallel, IMARA has also developed the “Wireless Infrastructure to Vehicle communication system”. Here, a specific hardware and software systems have been developed to allow V2I and V2V applications integrated on the Joint System demonstrator. This architecture – called the 4GCube – is based on the wireless communications devices and standards (802.11 a/b/g); they are IPv6 / CALM compliant architectures that have been tested with 2.4 GHz and 5 GHz bands, allowing multiple services handling.

The integration and the validation of the Co-pilot system as well as the communication device were done during the Final Event of HAVEit with a tremendous success and with high professional standards. The FASCar was able to demonstrate high driving autonomous skills at high speeds (up to 120 km/h) and was able to demonstrate new advanced features like overtaking mobile and static obstacles.

6.2. New urban transportation platforms: INRIA’s Cybus

Participants: Laurent Bouraoui, François Charlot, Carlos Holguin, Fawzi Nashashibi, Tony Noël, Michel Parent, Paulo Lopes Resende, Jianping Xie, Armand Yvet.
In order to achieve autonomous driving, autonomous systems (robots, intelligent vehicles, UAV's, UGV's,...) must have a decisional system that integrates an advanced perception system that performs sensors data fusion and environment modeling. From perception to control, task planning and path planning algorithms have to plan safe and optimized itineraries while processing sensory data. Motion control is the last link of the processing chain where itineraries and dynamic trajectories are executed by the low level control system. IMARA works on each of the topics mentioned above.

With the European projects CityMobil and CityNet Mobil, IMARA had the opportunity to validate the autonomous driving architecture developed through the projects showcases held in European cities. In 2011, several cities hosted IMARA team in charge of demonstrating autonomous driving and autonomous sensor-based navigation using sensory data (laser scanners) and GNSS (GPS). These events were a total success and were the opportunity to deploy and test the large-scale SLAMMOT system used for environment mapping, vehicle localization and mobile obstacles detection but also the new VMS (Vehicle Management System) developed in order to coordinate the mobility and navigation of several Cybercars (INRIA's Cybus platforms).

For this purpose, the perception sub-system was based on the generic SLAM-based system that was already presented in detail last year [43]. The proposed localization architecture has been implemented in two different vehicle platforms. AGV is a fully autonomous vehicle equipped with two IBEO Alasca-XT laser scanners (left and right front corner). Cycab is a prototype of smart car mounted with a single IBEO-ML laser scanner in front.

This year, because of delivery problems related to an industrial partner, IMARA had to design and develop its own cybercars. Thus the Cybus are the newest prototyping and demonstration platforms designed at INRIA. Apart from the chassis and engines, the whole hardware and software systems were developed thanks to IMARA’s researchers and engineers talents. These electric vehicles are based on a Yamaha chassis but the embedded intelligence is the result of this year’s IMARA developments. Much of the perception and control software is now registered.

The system developed can be seen as an experimental platform for a new public mobility transportation system operating in mixed environments. This electric vehicle is a “clean” transportation mean that is capable to achieve the well known last-mile itinerary in urban areas where classical transportation means are inefficient or simply non profitable. In order to demonstrate the feasibility of the system, a 3-months service has been programmed in the City of La Rochelle. The Mayor of the City authorized the Cybus to operate during 3-months providing free transport service to the inhabitants of the city. For that purpose, a 2 kilometers route has been defined on which 5 calling stations where installed. The users were able to call the Cybus from any station to reach any other station. The evaluation of the system has shown an acceptability of more than 95% among population; very low failures or technical leakages were reported. Following this successful operation,
the European Commission asked INRIA to extend this operation as well as its technical capabilities in order to achieve a simultaneous multiple Cybus navigation. This was achieved last December and was a real success.

The platforms developed here (Cybus) were exploited and demonstrated in the context of the EU-CityNetMobil project. The cities of Antibes (France) and Reggio di Calabria (Italy) hosted the team for 2 weeks respectively in order to experiment this new mobility and transportation mean in the heart of their cities.

Following this success, a new proposal of CityMobil-2 project is under submission with the objective this time to extend real operational mobility services to 6-12 months in selected European cities!

6.3. Communications Management in Cooperative Intelligent Transportation Systems

Participants: Thierry Ernzt, Manabu Tsukada.

Cooperative Intelligent Transportation Systems (Cooperative ITS) are systems where the vehicles, the roadside infrastructure, central control centers and other entities exchange information in order to achieve better road safety, traffic efficiency and comfort of the road users. This exchange of information must rely on a common communication architecture. The ITS Station reference architecture has thus been specified in ISO and ETSI. It allows vehicles and roadside ITS stations to organize themselves into Vehicular Ad-hoc Network (VANET), presumably though IPv6 GeoNetworking using IEEE802.11p and to connect seamlessly to the
Internet though any available access technology. Several paths may thus be available at a given vehicle ITS station to communicate with other ITS stations. Paths are of three types: direct path, optimize path and anchor path. The objective of the study is to optimize the communication between ITS Stations by selecting the best available communication path. This requires first to gather information available locally at the ITS station (position, speed, application requirements, media characteristics, capabilities, path status, ...) and collected from neighbors ITS stations (position, speed, services, ...) and then to process this information through a decision-making algorithm. First, we define a network module allowing the combination of IPv6 together with GeoNet-working. Second, we propose a cross-layer path selection management module. Our contributions are mapped to the ITS station reference architecture by defining the relation between the ITS station network and transport layer (which hosts our IPv6 GeoNetworking contribution) and the vertical ITS station cross-layer entity (which hosts the path decision-making algorithm). We specify the functions allowing the exchange of parameters through the Service Access Point (SAP) between the network layer and the management entity (MN-SAP). The parameters used at the cross layer ITS station management entity are abstracted in a way so that they are agnostic to the protocols used at the ITS station network and transport layer, therefrom allowing easy replacement of protocol elements (e.g. replacing NEMO by other mobility support protocol) or permutation of the network stack (IPv6 or GeoNetworking, a combination of both or other network stack).

6.4. Managing the system (via probabilistic modeling)

6.4.1. Belief propagation inference for traffic prediction

Participants: Cyril Furtlehner, Jean-Marc Lasgouttes, Arnaud Lewden, Victorin Martin.

This work 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 are done in particular in the framework of the projects TRAVESTI and Pumas.

This year’s highlights are

- A particular effort has been done this year in studying the theoretical aspects of the ways to encode real valued variable into an binary Ising model. A publication on the subject is in preparation.
- A review of our work on road traffic inference using methods from statistical physics has been published.
- The investigation of the effect of various types of normalization in the belief propagation algorithm has lead to a technical report.
- Arnaud Lewden has specified and implemented the new software BPstruction, which is our contribution to the Pumas project. Besides implementing traffic reconstruction from FCD, it is intended as a testbench for our research on inference using Belief Propagation.
- Victorin Martin has given a talk at the “Séminaire de Modélisation des Réseaux de Transport” at IFSTTAR. He presented there our method for real-time traffic reconstruction and prediction.
- Jean-Marc Lasgouttes also presented this work at the Xerox Research Centre Europe seminar.

6.4.2. Evaluation of dual mode transport system by event-driven simulation

Participants: Arnaud de La Fortelle, Sami Mahari.

The European project CATS — City Alternative Transport System — is developing and evaluating a new vehicle system using a single type of vehicle for two different usages: individual use or collective transport. Real experiments will necessarily take place with a limited number of vehicles and stations. Hence there is a need for evaluation using simulations. We have been developing a discrete events simulator for that purpose, based on a previous work done for collective taxis.
Our model relies on an adapted events/decision graph that extends previous graphs. The new feature of this model is the way we deal with two modes that can be extended to many other modes. This work therefore shows on a concrete example a method to efficiently merge multiple modes into one model.

- This year has seen the design and first implementation of the simulator.
- The results have been presented at a conference [29].

6.4.3. Multi-speed exclusion processes

**Participants:** Cyril Furtlehner, Jean-Marc Lasgouttes, Maxim Samsonov.

The slow-to-start mechanism is known to play an important role in the particular shape of the Fundamental diagram of traffic and to be associated to hysteresis effects of traffic flow. We study this question in the context of stochastic processes, namely exclusion and queueing processes, by including explicitly an asymmetry between deceleration and acceleration in their formulation. Spatial condensation phenomena and metastability are observed, depending on the level of the aforementioned asymmetry. The relationship between these 2 families of models is analyzed on the ring geometry, to yield a large deviation formulation of the fundamental diagram (FD)

This work has been presented at the TGF’11 conference [22], and a more extensive article is in preparation for a journal.

6.4.4. Dynamics of points of interest in a social game

**Participants:** Guy Fayolle, Jean-Marc Lasgouttes.

*Ma Micro Planète* is a geolocalized video game which entices players to use sustainable means of transport. 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 examine the system in the thermodynamic limit, as the number of players tends to infinity, the existence of which is proved under general conditions, where the probability of increasing the score of a visited POI is a function of the state of the system. Concerning the existence of a stationary regime, some complete answers are given for particular values of the parameters, and the existence of possible phase transition phenomena is enlightened.

A publication on the subject is in preparation.

6.4.5. Random walks in the quarter plane

**Participant:** Guy Fayolle.

In collaboration with K. Raschel (CNRS, Université F. Rabelais à Tours), we pursued the works initiated in 2010 in two main directions.

6.4.5.1. The zero drift case

In several recent studies on random walks with small jumps in the quarter plane, it has been noticed that the so-called group of the walk governs the behavior of a number of quantities, in particular through its order. In the article [11], when the drift of the random walk is equal to 0, we provide an effective criterion giving the order of this group. More generally, we also show that in all cases where the genus of the algebraic curve defined by the kernel is 0, the group is infinite, except precisely for the zero drift case, where finiteness is quite possible.

6.4.5.2. Counting and asymptotics

The enumeration of planar lattice walks, is a classical topic in combinatorics. For a given set $S$ of allowed unit jumps (or steps), it is a matter of counting the number of paths starting from some point and ending at some arbitrary point in a given time, and possibly restricted to some regions of the plane.

Like in the probabilistic context, a common way of attacking these problems relies on the following analytic approach. Let $f(i, j, k)$ denote the number of paths in $\mathbb{Z}_2^+$ starting from $(0, 0)$ and ending at $(i, j)$ at time $k$. Then the corresponding CGF
\[ F(x, y, z) = \sum_{i,j,k \geq 0} f(i, j, k)x^iy^jz^k \]

satisfies the functional equation

\[ K(x, y)F(x, y, z) = c(x)F(x, 0, z) + \tilde{c}(y)F(0, y, z) + c_0(x, y), \]

where \( x, y, z \) are complex variables, although the time variable \( z \) plays somehow the role of a parameter. The question of the type of the associated counting generating functions, that is rational, algebraic, holonomic (solution of a linear differential equation with polynomial coefficients), was solved whenever the group is finite (see RA 2010). When the group is infinite, the problem is still largely.

It turns out that the nature of singularities play a deep important role in this classification. Making use of the general and powerful approach proposed in the book [3], a paper entitled Some exact asymptotics in the counting of walks in the quarter-plane has been submitted to AofA (International Conference on Analysis of Algorithms, Montreal, June 2012), in which a new approach is proposed to obtain some exact asymptotics for walks confined to the quarter plane.

6.4.6. Statistical physics and hydrodynamic limits

Participants: Guy Fayolle, Cyril Furtlehner.

Having in mind a global project concerning the analysis of complex systems, we first focus on the interplay between discrete and continuous description: in some cases, this recurrent question can be addressed quite rigorously via probabilistic methods.

To attack this class of problems, in touch with many applications domains (e.g. biology, telecommunications, transportation systems), we started from paradigmatic elements, namely the discrete curves subjected to stochastic deformations, as those mentioned for instance in [39]. After convenient mappings, it appears that most models can be set in terms of interacting exclusion processes, the ultimate goal being to derive hydrodynamic limits for these systems after proper scalings. We extend the key ideas of [39], where the basic ASEP system on the torus was the toy model. The usual sequence of empirical measures, converges in probability to a deterministic measure, which is the unique weak solution of a Cauchy problem.

The Gordian knot is the analysis of a family of specific partial differential operators in infinite dimension. Indeed, the values of functions at given points play here the role of usual variables, their number becoming infinite. The method presents some new theoretical features, involving promeasures (as introduced by Bourbaki), variational calculus, functional integration, and the construction of generalized measures. In [20], we present a detailed analysis of the ASEP system on the torus \( \mathbb{Z}/N\mathbb{Z} \). Then we claim that most of the arguments a priori work in higher dimensions (ABC, multi-type exclusion processes, etc), leading to systems of coupled partial differential equations of Burgers’ type. In the course of the study, several fascinating multi-scale problems emerge quite naturally, bringing to light some connections with the so-called renormalization in theoretical physics.
6. New Results

6.1. Feature space modelling

6.1.1. A novel shape boundary based description for leaf identification

Participants: Itheri Yahiaoui, Olfa Mzoughi, Nozha Boujemaa.

The problem of automatic leaf identification is particularly difficult for two main reasons: (i) the first is the enormous number of leaf species and (ii) the second, which is relevant for some special species but more complex, is the high inter-species and the low intra-species similarity.

Our research has focused on analysing leaf morphology in order to determine a numeric key description for leaf species robust to all the above mentioned constraints. The approach that we propose is a shape boundary description that combines two complementary information: (i) the first one outlines local variations of the leaf margin. This is performed using the Directional Fragment Histogram (DFH), introduced in [25], which encodes the relative frequency distribution of groups of contour points with uniform orientation, (ii) the second property emphasizes the spatial distribution of contour points (in terms of distances). This is done by comparing the shape to standard geometric ones (such as circle, rectangle, ellipse, convex hull, etc.).

This descriptor was evaluated within the framework of ImageCLEF 2011 plant task where a crowd-sourced database, called Pl@ntLeaves [13], was used and a high number of image retrieval techniques was tested (a total of 8 groups from all around the world that have submitted 20 runs [19]). Our descriptor brought the best rate for scan-like pictures and was close to the best rate for scan pictures. Besides to the accuracy, this descriptor requires very low computational time, which accomplishes a basic condition for real world application.

6.1.2. Visual-based plant species identification from crowdsourced data


Inspired by citizen sciences, the main goal of this work is to speed up the collection and integration of raw botanical observation data, while providing to potential users an easy and efficient access to this botanical knowledge. We therefore designed and developed an original crowdsourcing web application dedicated to the access of botanical knowledge through automated identification of plant species by visual content.

Technically, the first side of the application deals with content-based identification of plant leaves. Whereas state-of-the-art methods addressing this objective are mostly based on leaf segmentation and boundary shape features, we developed a new approach based on local features and large-scale matching. This approach obtained the best results within one sub-task of ImageCLEF 2011 plant identification benchmark [19]. The second side of the application deals with interactive tagging and allow any user to validate or correct the automatic determinations returned by the system.

Overall, this collaborative system allows to enrich automatically and continuously the visual botanical knowledge and therefore to increase progressively the accuracy of the automated identification. A demo of the developed application was presented at ACM Multimedia conference [13]. This work was done in collaboration with INRIA team ZENITH and with the botanists of the AMAP UMR team (CIRAD). It is also closely related to a citizen science project around plant’s identification that we developed with the support of the TelaBotanica social network inside the Pl@ntNet project.

6.1.3. Spatial relations between salient points on a leaf

Participants: Sofiène Mouine, Itheri Yahiaoui, Anne Verroust-Blondet.
Figure 1. Normalized classification scores for scan images (up) and scan-like images (down). Our approach bin is in red, other ImageCLEF methods in blue.
In the scope of the Pl@ntNet project, our recent work has consisted in finding spatial relationships between salient points on a leaf. As a first step, classic detectors were used to find significant points in the leaf area and then the Shape context descriptor, originally applied on contour points, was introduced to measure a spatial relation between interest points. We have tested different configurations by varying the set of voting points. First results confirm that including spatial relations enriches the local description of each point. We are currently improving a veins and landmark extraction approach [12] in order to include also veins points in the voting set.

6.1.4. 3D mesh segmentation by example

Participants: Esma Elghoul, Anne Verroust-Blondet.

In recent years, there has been an increasing interest for automatic 3D segmentation. Indeed, segmentation of 3D objects is an important step in many applications such as part indexing of 3D objects, pattern recognition, compression, morphing, texture mapping and simplification. It refers to the process of partitioning 3D shapes into multiple parts, based on semantic criteria and/or geometric criteria.

Our work consists in introducing an approach to segment a 3D object class referring to a given segmented object from this class (we called it segmentation by example). The considered segmentation method is not automatic: we want to use interactive tools that proved advantageous to segment a 3D shape into relevant parts.

As a first task, we reviewed the state of the art in 3D segmentation techniques recently proposed in the literature. The different techniques were evaluated and classified for the purpose of choosing the more appropriate one for our work. We opted for extending the technique of random walks [24] to build an interactive tool of 3D segmentation.

For the second task, we had to solve a basic problem: that of similarly direct objects belonging to the same class. Indeed, each 3D model is provided in a random orientation in the space. In order to align objects of a same class, we used the alignment approach developed in [22] which computes 3 alignment axes. To properly orient our objects between them, we had to develop an additional process to the last one. It combines a 2D ICP and a 3D ICP approaches to give the best orientation among 48 possibilities and pair each two objects meshes.

So having a user-supplied already segmented model (model (1)) and a second model (model (2)) belonging to the same class (not segmented but similarly oriented), we developed a method to put into correspondence segmented parts of model (1) with faces of model (2). Then we computed a segmentation of model (2) using a derivative approach of the random walks. We applied this technique as well to segment all the objects that belong to the class of model (1). Our approach provides good results for manufactured object classes such as chairs and tables.

6.2. Feature space structuring

6.2.1. Random Maximum Margin Hashing

Participant: Alexis Joly.

Following the success of hashing methods for multidimensional indexing, more and more works are interested in embedding visual feature space in compact hash codes. Such approaches are not an alternative to using index structures but a complementary way to reduce both the memory usage and the distance computation cost. Several data dependent hash functions have notably been proposed to closely fit data distribution and provide better selectivity than usual random projections such as LSH. However, improvements occur only for relatively small hash code sizes up to 64 or 128 bits due to the lack of independence between the produced hash functions. In this work, we introduced a new hash function family that attempts to solve this issue in any kernel space. Rather than boosting the collision probability of close points, this method focus on data scattering. By training purely random splits of the data, regardless the closeness of the training samples, it is indeed possible to generate consistently more independent hash functions. On the other side, the use of large margin classifiers allows to maintain good generalization performances. Experiments did show that our
new Random Maximum Margin Hashing scheme (RMMH) outperforms four state-of-the-art hashing methods, notably in kernel spaces. Overall, this new concept of randomly trained classifiers opens the door to many other problems including large-scale learning, visual vocabulary construction or distributed content-based retrieval methods. A paper describing RMMH was published in the proceedings of CVPR 2011 [14].

6.2.2. Scalable information retrieval in distributed architectures

Participants: Mohamed Riadh Trad, Alexis Joly, Nozha Boujemaa.

Organizing media according to the occurrence of real-life events is attracting increasing interest in the multimedia community. However, whereas text based methods are now mature enough to deal with huge datasets, there are still some challenging issues managing multimedia contents. This becomes even more challenging in the context of User Generated Contents. Low-level visual metadata are indeed not simple textual or scalar values, their management requires efficient similarity search in high dimensional spaces.

Similarity search in high dimensional spaces has been the focus of many works in the database community in the recent years. State-of-the-art methods focus mainly on space partitioning techniques and more recently on hash-based probabilistic algorithms.

Although, hash-based approaches proved to be scalable, the computational cost is still too high for some real world applications and K-Nearest Neighbours Graph constructed can be more desirable than the costly online K-NN search. In fact, the basic LSH algorithm partitions the space uniformly and thus it does not exploit the clustering property of the data, which may result in slow query response and wasted space with additional hash tables. These limitations were pointed out with our scalable prototype for large scale event matching [18].

Scaling up LSH-based techniques and applications is then closely related to buckets occupations and objects distribution within the index structure. Recent works achieve better data distribution over the buckets with guarantees on occupation. As one result, we easily bound the similarity join size and evaluate bound algorithms complexity.

Based on these works, we designed and implemented a scalable prototype for distributed similarity search and K-NN graph construction. We have made several experiments querying real world large datasets. The prototype proved to be efficient for both search and K-NN graph construction.

Ongoing experiments process a 1.2 million images dataset. Results will be submitted for publication.

6.2.3. Visual similarity sensitive hashing methods for semantic image search in very large collections of images


With the rapid development of information acquisition technology, we have witnessed an explosive growth in the scale of shared data collections. Then, it is now possible to tackle fundamental problems with very large datasets’ context. Especially those addressing challenging tasks in machine learning for developing large scale approaches for multimedia retrieval and mining. Computer Vision is experiencing this paradigm shift, with large annotated image and video datasets becoming available. Indeed, various benchmark datasets for image classification have been released such as image-net and LabelMe. Therefore, a key challenge is taken up through out the Phd aiming to build efficient methods for training and matching efficiently very large collections of images.

We proposed several SVM-based strategies to build new supervised hash function families from large annotated collections of features. We indeed investigated with an approach consisting in benefiting from different embedding approaches in order to build compact codes indexed with efficient similarity search structures. Therefore, we have extended a kernelized hashing method [14] with multi-class SVM to solve a K-class classification problem by choosing the maximum applied to the outputs of K SVMs. We indeed proposed hashing methods based on the multi-class SVM classification strategies: One vs One (OVO) And One vs All (OVA). An important task during this process was to experimentally evaluate the quality lost induced by such representations with respect to the efficiency gains. We then compared multi-class SVM strategies with different underlying kernels.
Inspired by state of the art hashing in kernel space methods we investigated an approach consisting of benefiting from both semantic hashing like techniques and kernel embedding approach in order to build compact category aware codes indexed with efficient similarity. Experiments, are performed on image-net ILSVRC 2010 dataset [23]. Results will be submitted for publication.

6.3. Pattern recognition and statistical learning

6.3.1. Machine identification of biological shapes

Participants: Asma Rejeb Sfar, Donald Geman, Nozha Boujemaa.

Stored images of biological objects are accumulating at a staggering rate due to new sensor technologies, expanding use in medical diagnostics, web-based search engines and growing demands for web-based services in traditional sciences such as botany. These developments have been accompanied by an increasing demand for the automated analysis of these data, such as counting cell types, detecting lesions and other abnormalities in medical images, and identifying botanical shapes.

All these tasks have one feature in common: massive diversity among the shapes. Indeed, such shapes display enormous within-class variation and are generally highly deformable. Also, they often exhibit a hierarchical organization resulting from evolutionary processes.

There is currently no existing methodology in image analysis and computer vision which can be applied to a multi-class shape recognition problem of this complexity. Consequently, there is a need for a new, generic methodology for categorizing hierarchically-structured families of deformable shapes, particularly when both the number of categories and the within-category variation are very large.

We proposed a coarse-to-fine (CTF) approach in both shape representation and image parsing. The representation is hierarchical in both class and pose.

We focused on botanical shapes, specifically categorizing simple leaves according to species. So, we determined a suitable representation for the pose of a simple leaf and designed and tested a two-stage pose detector. Then, we constructed classifiers based on the plant taxonomy.

Results will be submitted for publication.

6.4. Interactive search and personalisation

6.4.1. Database denoising and multi visual queries

Participant: Sébastien Poullot.

One of IMEDIA’s task inside the SCARFACE project is to introduce and develop a character retrieval system. For this purpose, we take as entries the tracking of the persons in video sequences computed by Thalès and construct a database of the profiles. A profile is a 3D frame, a bounding box that changes along the time line. Two original works have been proposed for searching in the profile database.

The first one consists in analysing features of each profile with respect to all the profiles in order to extract relevant features from it, and construct more representative databases.

The second one is to be able to search inside the database with a set of queries (pictures of the same person). An a priori work can be done on this set of queries in order to extract the relevant features (and remove the irrelevant ones). On the other side an a posteriori work can be done on late merging depending on the specificities of each sub query.

TRECVID Instance Search 2011

Before starting the developments for SCARFACE, we tested various algorithms in TRECVID 2011 INS (instance search) task. This task is close to the SCARFACE one: from a set of captures of one object, one should find its occurrences in a set of video sequences. This work has been done during the stay of Sébastien Poullot at NII (the Japan National Institute of Informatics) in July and August 2011.
The differences with SCARFACE are:

- a high diversity in the type of the objects (people but also, places, vehicle, animals, etc),
- the location of the object in the database is not given.

Our approach obtains good results (above the median scores of all teams) and works in a very short time (and without indexing system for speeding up the process) [15]. The choice for SCARFACE’s method partially depends on these results. We still continue on the INS task in order to achieve better scores (various descriptors and various post and late fusion between sub queries).

**Query generative models**

Moreover, in order to enhance visual query results, we want to create some visual query generative models. It is directly linked to SCARFACE (a priori processes) and TRECVID works: given a set of images (considered as queries), we extract what gather them and what separate them in order to construct artificial relevant queries.

For now we essentially work on some logo databases.

### 6.4.2. Object-based Visual Query Suggestion

**Participants:** Amel Hamzaoui, Pierre Letessier, Alexis Joly, Nozha Boujemaa.

After our work on the shared neighbours clustering methods in multi-sources case published in [10], we are interested now to the case of a bipartite graph that we apply to object-based visual query suggestion using the visual words mining technique [16]. In fact, state-of-the-art visual search systems allow to retrieve efficiently small rigid objects in very large datasets. They are usually based on the query-by-window paradigm: a user selects any image region containing an object of interest and the system returns a ranked list of images that are likely to contain other instances of the query object. User’s perception of these tools is however affected by the fact that many submitted queries actually return nothing or only junk results (complex non-rigid objects, higher-level visual concepts, etc.). We address the problem of suggesting only the object’s queries that actually contain relevant matches in the dataset. This requires to first discover accurate object’s clusters in the dataset (as an off-line process); and then to select the most relevant objects according to user’s intent (as an online process). We therefore introduce a new object’s instances clustering framework based on two main contributions: efficient object’s seeds discovery with adaptive weighted sampling and bipartite shared-neighbours clustering. Experiments show that this new method outperforms state-of-the-art object mining and retrieval results on OxfordBuilding dataset. We finally describe two object-based visual query suggestion scenarios using the proposed framework and show examples of suggested object queries.

### 6.4.3. Interpretable Visual Models for Human Perception-based Object Retrieval

**Participants:** Ahmed Rebai, Alexis Joly, Nozha Boujemaa.

Understanding the results returned by automatic visual concept detectors is often a tricky task making users uncomfortable with these technologies. In this work we attempt to build humanly interpretable visual models, allowing the user to visually understand the underlying semantic. We therefore proposed a supervised multiple instance learning algorithm that selects as few as possible discriminant local features for a given object category. The method finds its roots in the lasso theory where a L1-regularization term is introduced in order to constraint the loss function, and subsequently produce sparser solutions. Efficient resolution of the lasso path is achieved through a boosting-like procedure inspired by Blasso algorithm. Quantitatively, the method achieved similar performance as current state-of-the-art, and qualitatively, it allows users to construct their own model from the original set of patches learned, thus allowing for more compound semantic queries. This work is part of the PhD of Ahmed Rebai [8] and it was published in ICMR 2011 proceedings [17]. This work was then extended to using geometrically checked feature sets rather than using single local features to describe the content of visual patches. We did show that this allows drastically reducing the number of the selected visual words while improving their interpretability. A publication was submitted to pattern recognition journal [11].
6.4.4. Relevance feedback on local features: Application to plants annotations and identification

Participants: Wajih Ouertani, Michel Crucianu, Nozha Boujemaa.

As biological image databases are increasing rapidly, automated species identification based on digital data is of great interest for accelerating biodiversity assessment, researches and monitoring. In this context, our work falls within an investigation of computer vision techniques or more precisely: object recognition and content based image retrieval techniques to help botanist identifying and organizing his digital images’ collections. Under believe that perception, recognition and decision are parts of human skills, this work focus on an interactive mechanism which tries to extract useful information from the user and gives him help to deal with large data amount. We adopted an explicit relevance feedback (RF) schema and we worked on extending it to deal with local intention through local features (LF) description. This mechanism helps discovering and dynamically defining new concept and interesting plant parts and feed identification ways interactively. Moreover since it relies to the content rather than labels one direct application is to fill the initially sparse annotation space with right annotations and in a reasonable time and with the introduce of one or many expertises. We recently explored and tested images local features matching involving high order features and non-rigid adaptation tentative to structure database with a patterns’ discovery stage. Using those type of methods we expect to introduce a high level appearance information that tends to go beyond classical bag of features and histogram based distances at least from semantic gap and interpretation point of view. We argue our exploration way with the fact that initial search space can be exceedingly rich. By pre-structuring it we can hope to obtain a smaller search space together with more reliable inference. Also learning parts interactively with localized local features may require a lot of interaction since it requires a considerable number of examples. We experienced the design of combined machine learning and prior mining of matches which we are actually improving.

6.5. Software

6.5.1. IKONA/MAESTRO software


This year, IKONA has been extended in the context of Pl@ntNet, Glocal, I-SEARCH and R2I projects. For the Pl@ntNet project, along the continuing improvements in the MAESTRO software, a number of new features were added. Namely the support for the automatic image segmentation and subsequent use of segmented regions; descriptors with the various shape’s geometric parameters; use of multiple orientations for Harris points; run-time additions to the external database and immediate availability of the new images for the search; descriptors to facilitate external data usage; colour SIFT and Affine Covariance descriptors; integration of the thesis work of Ahmed Rebai for objects retrieval; and tools for statistical tests.

In addition, a number of new web services were developed and deployed: the dynamic indexation system of the on-line pad (“carnet-en-ligne” of Tela Botanica)images; the search with multiple views; the update of Pl@ntNet internal demonstration allowing to present features such as visual similarity search, textual search, filtering (pre- and post-filtering), and different methods of research; the implementation of the organ prediction web service and other web services of botanical information statistics; the administration of the indexation system and the experimentation of new research methods (GPS spatial and temporal search).

For the Glocal project, an interface was developed for the demonstrations of a search engine in large scale events database (the queries are event images and the result is a list of the closest events in terms of time alignment and image content), and new web services were developed and updated according to the data exchange format and the middle-ware of the project - among others: fraud detection, import media from the web, associate media with existing event in the repository, and event matching web services. The queries are composed of either a medium link (an external image) or an event link (set of external images).
For the I-SEARCH project, an integration was performed to provide global and local 2D image low level descriptors. For videos, an automated extraction of visual words tool was integrated to show to users image patches which are the most meaningful.

For the R2I project, a detailed technical documentation of the procedure of maestro’s installation, web services and tomcat server were provided to Exalead partner.
6. New Results

6.1. Embedded data management

Participants: Nicolas Anciaux, Luc Bouganim, Yanli Guo, Lionel Le Folgoc, Philippe Pucheral, Shaoyi Yin.

Inspired by low cost economic models, this work draws the idea of a one-dollar database machine, with the objective to disseminate databases everywhere, up to the lightest smart objects. In contrast to traditional database machines relying on massively parallel architectures, the one-dollar database machine considers the cheapest form of computer available today: a microcontroller equipped with GBs size (external) Flash storage. Designing such a database machine is very challenging due to a combination of conflicting RAM and NAND Flash constraints. To tackle this challenge, this work proposes a new paradigm based on database serialization (managing all database structures in a pure sequential way) and stratification (restructuring them into strata when a scalability limit is reached). We show that a complete DBMS engine can be designed according to this paradigm and demonstrate the effectiveness of the approach through a performance evaluation.

This work capitalizes on previous results related to the indexing of Flash resident data [39] and has also obvious connections with the more general study we are conducting on Flash-based data management (see Section 6.2). Partial elements of this solution have been demonstrated at [28].

6.2. Flash-based Data Management

Participant: Luc Bouganim.

Bimodal flash devices. While disks have offered a stable behavior for decades, thus guaranteeing the timelessness of many database design decisions, flash devices keep on mutating. Many researchers have proposed to adapt database algorithms to existing flash devices. However, today, there is no reference DBMS design based on solid assumptions of flash devices behavior, precisely because flash device behavior varies across models, across firmware updates and possibly over time for the same model: database researchers are running after flash memory technology. In this study, we took the reverse approach and defined how flash devices should support database management. We advocated that flash devices should provide guarantees to a DBMS so that it can devise stable and efficient IO management mechanisms. Based on the characteristics of flash chips, we defined a bimodal FTL that distinguishes between a minimal mode where sequential writes, sequential reads and random reads are optimal while updates and random writes are forbidden, and a mode where updates and random writes are supported at the cost of sub-optimal IO performance. This work started at the end of 2010 and was published at CIDR’11 [19], in cooperation with the IT University of Copenhagen. DBMS/Flash device co-design considerations were the focus of a tutorial on flash devices given recently at VLDB 2011 [20].

6.3. Privacy-Preserving Data Publishing

Participants: Tristan Allard, Benjamin Nguyen, Philippe Pucheral.

While most PPDP works make the assumption of a trusted central publisher, this study advocates a decentralized way of publishing anonymized datasets. More precisely, our work concerns the proof of feasibility of adapting traditional PPDP schemes, such as $k$-anonymity, $\ell$-diversity or differential privacy to encompass the use of secure portable devices. In the applications we consider, each secure device is a data provider with weak computing capacities and weak connectivity (frequency and duration of connections are unpredictable).1 E.g., in the e-health context, patients may have their medical folder embedded in a secure device and connect it sporadically when they visit their physician or when they want to consult it at home.

---

1 E.g., in the e-health context, patients may have their medical folder embedded in a secure device and connect it sporadically when they visit their physician or when they want to consult it at home.
connectivity precludes any P2P solution to the problem. A server allowing asynchronous communications between the devices becomes necessary to implement a distributed PPDP mechanism but this server does not benefit from the same trustworthiness as the participating devices. Our work aims to provide a generic method to adapt an important subclass of PPDP algorithms to this context, using both the limited secure computation capacities of each device (but taking advantage of their number) and the powerful computation abilities of an untrusted server available 24/7. Our proposal is based on a meta algorithm divided in three phases: (1) a collection phase where encrypted data is collected by the untrusted server, (2) a construction phase where the untrusted server performs a sound computation of a given privacy mechanism to generate sanitization rules and (3) a sanitization phase where the encrypted data is decrypted then sanitized by the devices to produce a final clear-text result. The last phase can be distributed using many different devices for better efficiency. In [15], [17], we showed how it is possible to transform existing anonymity mechanisms into decentralized ones using secure devices, while maintaining equivalent security guarantees against honest-but-curious and weakly malicious adversaries. In [16], we studied the (unlikely) event that some secure devices might be compromised, and can collude with the untrusted server. We provided schemes to detect the compromised devices with a probability that can be fixed as close to 1 as desired (the trade-off being the latency of the protocol).

6.4. Minimal Exposure

**Participants:** Nicolas Anciaux, Benjamin Nguyen.

When users request a service, the service provider usually asks for personal documents to tailor its service to the specific situation of the applicant. For example, the rate and duration of consumer’s loans are usually adapted depending on the risk based on the income, assets or past lines of credits of the borrower. In practice, an excessive amount of personal data is collected and stored. Indeed, a paradox is at the root of this problem: service providers require users to expose data in order to determine whether that data is needed or not to achieve the purpose of the service. We currently explore a reverse approach, where service providers would publicly describe the data they require to complete their task, and where the applicants would confront those descriptions with their own data to determine themselves the minimal subset of information to expose. We have first investigated solutions for simplistic tasks (e.g., evaluating a decision tree to determine the loan rate and duration a given applicant can claim), and we plan to address more complex ones (e.g., building the profile of customers, mining association rules, etc.) in the short term. The work on Minimal Exposure has just started and a first paper is under evaluation.

6.5. Experiment in the medical field

**Participants:** Nicolas Anciaux, Luc Bouganim, Lionel Le Folgoc, Philippe Pucheral, Alexei Troussov.

The PlugDB engine is being experimented in the field since September 2011 to implement a secure and portable medical-social folder. The objective is to improve the coordination of medical care and social services provided at home for dependent people. Details related to this experiment conducted with about 120 practitioners and patients are given in Section 7.2. While this action did not generate new academic results (though it helped us validating some previous results), it imposed us a strong investment in terms of test and optimization for our prototype and in terms of communication to promote this experiment at the regional level.
WILLOW Project-Team

6. New Results

6.1. 3D object and scene modeling, analysis, and retrieval

6.1.1. Quantitative image analysis for archeology

Participants: Bryan Russell, Jean Ponce, Josef Sivic, Helene Dessales [ENS Archeology laboratory].

Figure 1. (a) Example photographs captured of the Pompeii site (563 photographs are used in total). (b) Rendered viewpoints of the recovered 3D model. Notice the fine-level details that are captured by the model.

Accurate indexing and alignment of images is an important problem in computer vision. A successful system would allow a user to retrieve images with similar content to a query image, along with any information associated with the image. Prior work has mostly focused on techniques to index and match photographs depicting particular instances of objects or scenes (e.g. famous landmarks, commercial product labels, etc.). This has allowed progress on tasks, such as the recovery of a 3D reconstruction of the depicted scene.

However, there are many types of images that cannot be accurately aligned. For instance, for many locations there are drawings and paintings made by artists that depict the scene. Matching and aligning photographs, paintings, and drawings is extremely difficult due to various distortions that can arise. Examples include perspective and caricature distortions, along with errors that arise due to the difficulty of drawing a scene by hand.

In this project, we seek to index and align a database of images, paintings, and drawings. The focus of our work is the Championnet house in the Roman ruins at Pompeii, Italy. Given an alignment of the images, paintings, and drawings, we wish to explore tasks that are of interest to archaeologists and curators who wish to study and preserve the site. Example applications include: (i) digitally restoring paintings on walls where the paintings have disappeared over time due to erosion, (ii) geometrically reasoning about the site over time through the drawings, (iii) indexing and searching patterns that exist throughout the site.
Figure 2. Final alignment between the paintings and 3D model. For each example, left: painting; middle: 3D model contours projected onto painting; right: synthesized viewpoint from 3D model using recovered camera parameters. For the examples in (a-c), note how the final alignment is close to the painting. Our system handles paintings that depict the 3D structure of the scene over time and span different artistic styles and mediums (e.g. water colors, cross-hatching, copies of originals on engravings). Notice how the site changes over time, with significant structural changes (e.g. the wall murals decay over time, the columns change). Example failure cases are shown in (d,e).
Recently, we have addressed the problem of automatically aligning historical architectural paintings with 3D models obtained using multi-view stereo technology from modern photographs. This is a challenging task because of the variations in appearance, geometry, color and texture due to environmental changes over time, the nonphotorealistic nature of architectural paintings, and differences in the viewpoints used by the painters and photographers. Our alignment procedure consists of two novel aspects: (i) we combine the gist descriptor with the view-synthesis/retrieval of Irschara et al. to obtain a coarse alignment of the painting to the 3D model, and (ii) we have developed an ICP-like viewpoint refinement procedure, where 3D surface orientation discontinuities (folds and creases) and view-dependent occlusion boundaries are rendered from the automatically obtained and noisy 3D model in a view-dependent manner and matched to gPB contours extracted from the paintings. We demonstrate the alignment of XIXth Century architectural watercolors of the Casa di Championnet in Pompeii with a 3D model constructed from modern photographs using the PMVS public-domain multi-view stereo software. Figure 1 shows some of the captured photographs and snapshots of the 3D reconstruction of the site. Notice that the 3D reconstruction captures much detail of the walls and structures. Example painting to 3D model alignments are shown in figure 2.

This work resulted in a workshop publication [16].

6.1.2. Visual localization by linear combination of image descriptors

**Participants:** Josef Sivic, Akihiko Torii [Tokyo Institute of Technology], Tomas Pajdla [CTU in Prague].

In this work, we seek to predict the GPS location of a query image given a database of images localized on a map with known GPS locations. The contributions of this work are three-fold: (1) we formulate the image-based localization problem as a regression on an image graph with images as nodes and edges connecting close-by images; (2) we design a novel image matching procedure, which computes similarity between the query and pairs of database images using edges of the graph and considering linear combinations of their feature vectors. This improves generalization to unseen viewpoints and illumination conditions, while reducing the database size; (3) we demonstrate that the query location can be predicted by interpolating locations of matched images in the graph without the costly estimation of multi-view geometry. We demonstrate benefits of the proposed image matching scheme on the standard Oxford building benchmark, and show localization results on a database of 8,999 panoramic Google Street View images of Pittsburgh.

This work resulted in a publication [18].

6.2. Category-level object and scene recognition

6.2.1. Task-Driven Dictionary Learning

**Participants:** Julien Mairal, Jean Ponce, Francis Bach [INRIA SIERRA].

Modeling data with linear combinations of a few elements from a learned dictionary has been the focus of much recent research in machine learning, neuroscience and signal processing. For signals such as natural images that admit such sparse representations, it is now well established that these models are well suited to restoration tasks. In this context, learning the dictionary amounts to solving a large-scale matrix factorization problem, which can be done efficiently with classical optimization tools. The same approach has also been used for learning features from data for other purposes, e.g., image classification, but tuning the dictionary in a supervised way for these tasks has proven to be more difficult. In this paper, we present a general formulation for supervised dictionary learning adapted to a wide variety of tasks, and present an efficient algorithm for solving the corresponding optimization problem. Experiments on handwritten digit classification, digital art identification, nonlinear inverse image problems, and compressed sensing demonstrate that our approach is effective in large-scale settings, and is well suited to supervised and semi-supervised classification, as well as regression tasks for data that admit sparse representations.

This work has resulted in a publication [4].

6.2.2. Ask the locals: multi-way local pooling for image recognition

**Participants:** Y-Lan Boureau, Jean Ponce, Nicolas Le Roux [INRIA SIERRA], Francis Bach [INRIA SIERRA], Yann LeCun [New York University].
Invariant representations in object recognition systems are generally obtained by pooling feature vectors over spatially local neighborhoods. But pooling is not local in the feature vector space, so that widely dissimilar features may be pooled together if they are in nearby locations. Recent approaches rely on sophisticated encoding methods and more specialized codebooks (or dictionaries), e.g., learned on subsets of descriptors which are close in feature space, to circumvent this problem. In this work, we argue that a common trait found in much recent work in image recognition or retrieval is that it leverages locality in feature space on top of purely spatial locality. We propose to apply this idea in its simplest form to an object recognition system based on the spatial pyramid framework, to increase the performance of small dictionaries with very little added engineering. State-of-the-art results on several object recognition benchmarks show the promise of this approach.

This work has resulted in a publication [7].

### 6.2.3. A Graph-matching Kernel for Object Categorization

**Participants:** Olivier Duchenne, Armand Joulin, Jean Ponce.

This paper addresses the problem of category-level image classification. The underlying image model is a graph whose nodes correspond to a dense set of regions, and edges reflect the underlying grid structure of the image and act as springs to guarantee the geometric consistency of nearby regions during matching. A fast approximate algorithm for matching the graphs associated with two images is presented. This algorithm is used to construct a kernel appropriate for SVM-based image classification, and experiments with the Caltech 101, Caltech 256, and Scenes datasets demonstrate performance that matches or exceeds the state of the art for methods using a single type of features.

This work has resulted in a ICCV 2011 publication [9] (oral presentation).

### 6.2.4. A Tensor-Based Algorithm for High-Order Graph Matching

**Participants:** Olivier Duchenne, Jean Ponce, Francis Bach [INRIA SIERRA], Inso Kweon [KAIST, Korea].

This paper addresses the problem of establishing correspondences between two sets of visual features using higher-order constraints instead of the unary or pairwise ones used in classical methods. Concretely, the corresponding hypergraph matching problem is formulated as the maximization of a multilinear objective function over all permutations of the features. This function is defined by a tensor representing the affinity between feature tuples. It is maximized using a generalization of spectral techniques where a relaxed problem is first solved by a multi-dimensional power method, and the solution is then projected onto the closest assignment matrix. The proposed approach has been implemented, and it is compared to state-of-the-art algorithms on both synthetic and real data.

This work has resulted in a PAMI publication [2].

### 6.2.5. Clusterpath: an algorithm for clustering using convex fusion penalties

**Participants:** Armand Joulin, Toby Hocking [INRIA SIERRA], Francis Bach [INRIA SIERRA], Jean-Philippe Vert [Mines ParisTech].

We present a new clustering algorithm by proposing a convex relaxation of hierarchical clustering, which results in a family of objective functions with a natural geometric interpretation. We give efficient algorithms for calculating the continuous regularization path of solutions, and discuss relative advantages of the parameters. Our method experimentally gives state-of-the-art results similar to spectral clustering for non-convex clusters, and has the added benefit of learning a tree structure from the data.

This work has resulted in an publication [10].

### 6.2.6. An MRF model for binarization of natural scene text

**Participants:** Karteek Alahari, Anand Mishra [IIT India], C.V. Jawahar [IIT India].
Scene text recognition has gained significant attention from the computer vision community in recent years. Recognizing text in the wild is a challenging problem, even more so than the recognition of scanned documents. In this work, we focus on the problem of cropped word recognition. We present a framework that exploits both bottom-up and top-down cues. The bottom-up cues are derived from individual character detections from the image. We build a Conditional Random Field model on these detections to jointly model the strength of the detections and the interactions between them. We impose top-down cues obtained from a lexicon-based prior, i.e. language statistics, on the model. The optimal word represented by the text image is obtained by minimizing the energy function corresponding to the random field model.

We show very significant improvements in accuracies on two challenging public datasets, namely Street View Text (over 15%) and ICDAR 2003 (over 10%).

This work has resulted in a publication [12].

6.2.7. Strongly-supervised deformable part model for object detection

Participants: Hossein Azizpour [KTH Stockholm], Ivan Laptev, Stefan Carlsson [KTH Stockholm].

Deformable part models achieve state-of-the-art performance for object detection while relying on the greedy initialization during training. The goal of this paper is to investigate limitations of such initialization and to improve the model for the case when part locations are known at the training time. To this end, we deploy part-level supervision and demonstrate improved detection results when learning models with manually-initialized part locations. We further explore the benefits of the strong supervision and learn model structure by minimizing the variance among adjacent model parts. Our method can simultaneously handle samples with and without part-level annotation making benefit even from a fraction of fully-annotated training samples. Experimental results are reported for the detection of six animal classes in PASCAL VOC 2007 and 2010 datasets. We demonstrate significantly improved performance of our model compared to the state-of-the-art LSVM object detector and poselet detector. Example learnt models are shown in figure 3.

This work has resulted in a submission to CVPR 2012.

6.2.8. Exploiting Photographic Style for Category-Level Image Classification by Generalizing the Spatial Pyramid

Participant: Jan van Gemert [University of Amsterdam].

This paper investigates the use of photographic style for category-level image classification. Specifically, we exploit the assumption that images within a category share a similar style defined by attributes such as colorfulness, lighting, depth of field, viewpoint and saliency. For these style attributes we create correspondences across images by a generalized spatial pyramid matching scheme. Where the spatial pyramid groups features spatially, we allow more general feature grouping and in this paper we focus on grouping images on photographic style. We evaluate our approach in an object classification task and investigate style differences between professional and amateur photographs. We show that a generalized pyramid with style-based attributes improves performance on the professional Corel and amateur Pascal VOC 2009 image datasets.

This work has resulted in a publication [20].

6.2.9. Generalized Fast Approximate Energy Minimization via Graph Cuts: Alpha-Expansion Beta-Shrink Moves

Participants: Karteek Alahari, Mark Schmidt [INRIA SIERRA].

We present alpha-expansion beta-shrink moves, a simple generalization of the widely-used alpha beta-swap and alpha-expansion algorithms for approximate energy minimization. We show that in a certain sense, these moves dominate both alpha beta-swap and alpha-expansion moves, but unlike previous generalizations the new moves require no additional assumptions and are still solvable in polynomial-time. We show promising experimental results with the new moves, which we believe could be used in any context where alpha-expansions are currently employed.
Figure 3. Illustration of deformable part models trained for the dog class. Left: dog detection with the proposed supervised part model. Right: dog detection with the original LSVM model of Felzenszwalb et al. Our model adapts to the different object appearance while LSVM model attempts to explain both samples using the same deformable HOG template.

This work has resulted in a publication [17].

6.3. Image restoration, manipulation and enhancement

6.3.1. Non-uniform Deblurring for Shaken Images

Participants: Oliver Whyte, Josef Sivic, Andrew Zisserman, Jean Ponce.

We argue that blur resulting from camera shake is mostly due to the 3D rotation of the camera, causing a blur that can be significantly non-uniform across the image. However, most current deblurring methods model the observed image as a convolution of a sharp image with a uniform blur kernel. We propose a new parametrized geometric model of the blurring process in terms of the rotational velocity of the camera during exposure. We apply this model in the context of two different algorithms for camera shake removal: the first uses a single blurry image (blind deblurring), while the second uses both a blurry image and a sharp but noisy image of the same scene. We show that our approach makes it possible to model and remove a wider class of blurs than previous approaches, and demonstrate its effectiveness with experiments on real images.

The project resulted in a publication [5].

6.3.2. Deblurring shaken and partially saturated images

Participants: Oliver Whyte, Josef Sivic, Andrew Zisserman.

We address the problem of deblurring images degraded by camera shake blur and saturated or over-exposed pixels. Saturated pixels are a problem for existing non-blind deblurring algorithms because they violate the assumption that the image formation process is linear, and often cause significant artifacts in deblurred outputs. We propose a forward model that includes sensor saturation, and use it to derive a deblurring algorithm properly treating saturated pixels. By using this forward model and reasoning about the causes of artifacts...
in the deblurred results, we obtain significantly better results than existing deblurring algorithms. Further we propose an efficient approximation of the forward model leading to a significant speed-up. Example result is shown in figure 4.

The project resulted in a publication [19].

6.3.3. Dictionary Learning for Deblurring and Digital Zoom

Participants: Florent Couzinie, Julien Mairal, Jean Ponce, Francis Bach [INRIA SIERRA].

This work proposes a novel approach to image deblurring and digital zooming using sparse local models of image appearance. These models, where small image patches are represented as linear combinations of a few elements drawn from some large set (dictionary) of candidates, have proven well adapted to several image restoration tasks. A key to their success has been to learn dictionaries adapted to the reconstruction of small image patches. In contrast, recent works have proposed instead to learn dictionaries which are not only adapted to data reconstruction, but also tuned for a specific task. We introduce here such an approach to deblurring and digital zoom, using pairs of blurry/sharp (or low-/high-resolution) images for training, as well as an effective stochastic gradient algorithm for solving the corresponding optimization task. Although this learning problem is not convex, once the dictionaries have been learned, the sharp/high-resolution image can be recovered via convex optimization at test time. Experiments with synthetic and real data demonstrate the effectiveness of the proposed approach, leading to state-of-the-art performance for non-blind image deblurring and digital zoom.

This work has resulted in a publication [1].

6.3.4. Sparse Image Representation with Epitomes

Participants: Louise Benoit, Julien Mairal, Jean Ponce, Francis Bach [INRIA SIERRA].

Sparse coding, which is the decomposition of a vector using only a few basis elements, is widely used in machine learning and image processing. The basis set, also called dictionary, is learned to adapt to specific data. This approach has proven to be very effective in many image processing tasks. Traditionally, the dictionary is an unstructured "flat" set of atoms. In this work, we study structured dictionaries which are
obtained from an epitome, or a set of epitomes. The epitome is itself a small image, and the atoms are all the patches of a chosen size inside this image. This considerably reduces the number of parameters to learn and provides sparse image decompositions with shift invariance properties. We propose a new formulation and an algorithm for learning the structured dictionaries associated with epitomes, and illustrate their use in image denoising tasks.

This work has resulted in a CVPR’11 publication [6].

6.3.5. Proximal Methods for Hierarchical Sparse Coding

Participants: Julien Mairal, Rodolphe Jenatton [INRIA SIERRA], Guillaume Obozinski [INRIA SIERRA], Francis Bach [INRIA SIERRA].

Sparse coding consists in representing signals as sparse linear combinations of atoms selected from a dictionary. We consider an extension of this framework where the atoms are further assumed to be embedded in a tree. This is achieved using a recently introduced tree-structured sparse regularization norm, which has proven useful in several applications. This norm leads to regularized problems that are difficult to optimize, and in this paper, we propose efficient algorithms for solving them. More precisely, we show that the proximal operator associated with this norm is computable exactly via a dual approach that can be viewed as the composition of elementary proximal operators. Our procedure has a complexity linear, or close to linear, in the number of atoms, and allows the use of accelerated gradient techniques to solve the tree-structured sparse approximation problem at the same computational cost as traditional ones using the l1-norm. Our method is efficient and scales gracefully to millions of variables, which we illustrate in two types of applications: first, we consider fixed hierarchical dictionaries of wavelets to denoise natural images. Then, we apply our optimization tools in the context of dictionary learning, where learned dictionary elements naturally self-organize in a prespecified arborescent structure, leading to better performance in reconstruction of natural image patches. When applied to text documents, our method learns hierarchies of topics, thus providing a competitive alternative to probabilistic topic models.

This work has resulted in a publication [3].

6.4. Human activity capture and classification

6.4.1. Track to the future: Spatio-temporal video segmentation with long-range motion cues

Participants: Jose Lezama, Karteek Alahari, Ivan Laptev, Josef Sivic.

Video provides rich visual cues such as motion and appearance but also much less explored long-range temporal interactions among objects. We aim to capture such interactions and to construct powerful intermediate-level video representation for subsequent recognition. Motivated by this goal, we seek to obtain spatio-temporal oversegmentation of the video into regions that respect object boundaries and, at the same time, associate object pixels over many video frames. The contributions of this paper are twofold. First, we develop an efficient spatio-temporal video segmentation algorithm, that naturally incorporates long-range motion cues from the past and future frames in the form of clusters of point tracks with coherent motion. Second, we devise a new track clustering cost-function that includes occlusion reasoning, in the form of depth ordering constraints, as well as motion similarity along the tracks. We evaluate the proposed approach on a challenging set of video sequences of office scenes from feature length movies.

This work resulted in a publication [11].

6.4.2. Density-aware person detection and tracking in crowds

Participants: Mikel Rodriguez, Ivan Laptev, Josef Sivic, Jean-Yves Audibert [INRIA SIERRA].
We address the problem of person detection and tracking in crowded video scenes. While the detection of individual objects has been improved significantly over the recent years, crowd scenes remain particularly challenging for the detection and tracking tasks due to heavy occlusions, high person densities and significant variation in people’s appearance. To address these challenges, we propose to leverage information on the global structure of the scene and to resolve all detections jointly. In particular, we explore constraints imposed by the crowd density and formulate person detection as the optimization of a joint energy function combining crowd density estimation and the localization of individual people. We demonstrate how the optimization of such an energy function significantly improves person detection and tracking in crowds. We validate our approach on a challenging video dataset of crowded scenes. The proposed approach is illustrated in figure 5.

This work has resulted in a publication [14].

![Figure 5. Individual head detections provided by state-of-the-art object detector (Felzenswalb et al. 2009) (bottom-left; green: true positives; red: false positives) are improved significantly by our method (bottom-right; yellow: new true positives) using the crowd density estimate (topright) obtained from the original frame (top-left).](image)

### 6.4.3. Data-driven Crowd Analysis in Videos

**Participants:** Mikel Rodriguez, Josef Sivic, Ivan Laptev, Jean-Yves Audibert [INRIA SIERRA].

In this work we present a new crowd analysis algorithm powered by behavior priors that are learned on a large database of crowd videos gathered from the Internet. The algorithm works by first learning a set of crowd behavior priors off-line. During testing, crowd patches are matched to the database and behavior priors are transferred. We adhere to the insight that despite the fact that the entire space of possible crowd behaviors is infinite, the space of distinguishable crowd motion patterns may not be all that large. For many individuals in a crowd, we are able to find analogous crowd patches in our database which contain similar patterns of behavior that can effectively act as priors to constrain the difficult task of tracking an individual in a crowd. Our algorithm is data-driven and, unlike some crowd characterization methods, does not require us to have seen the test video beforehand. It performs like state-of-the-art methods for tracking people having common crowd behaviors and outperforms the methods when the tracked individual behaves in an unusual way.

This work has resulted in a publication [15].

### 6.4.4. Learning person-object interactions for action recognition in still images

**Participants:** Vincent Delaitre, Josef Sivic, Ivan Laptev.
In this work, we investigate a discriminatively trained model of person-object interactions for recognizing common human actions in still images. We build on the locally order-less spatial pyramid bag-of-features model, which was shown to perform extremely well on a range of object, scene and human action recognition tasks. We introduce three principal contributions. First, we replace the standard quantized local HOG/SIFT features with stronger discriminatively trained body part and object detectors. Second, we introduce new person-object interaction features based on spatial co-occurrences of individual body parts and objects. Third, we address the combinatorial problem of a large number of possible interaction pairs and propose a discriminative selection procedure using a linear support vector machine (SVM) with a sparsity inducing regularizer. Learning of action-specific body part and object interactions bypasses the difficult problem of estimating the complete human body pose configuration. Benefits of the proposed model are shown on human action recognition in consumer photographs, outperforming the strong bag-of-features baseline. The proposed model is illustrated in figure 6.

This work has resulted in a publication [8].

Figure 6. Representing person-object interactions by pairs of body part (cyan) and object (blue) detectors. To get a strong interaction response, the pair of detectors (here visualized at positions $p_i$ and $p_j$) must fire in a particular relative 3D scale-space displacement (given by the vector $v$) with a scale-space displacement uncertainty (deformation cost) given by diagonal 3x3 covariance matrix $C$ (the spatial part of $C$ is visualized as a yellow dotted ellipse). Our image representation is defined by the max-pooling of interaction responses over the whole image, solved efficiently by the distance transform.

6.4.5. People Watching: Human Actions as a Cue for Single View Geometry

Participants: David Fouhey [CMU], Vincent Delaitre, Abhinav Gupta [CMU], Ivan Laptev, Alexei Efros [CMU], Josef Sivic.

We present an approach which exploits the coupling between human actions and scene geometry. We investigate the use of human pose as a cue for single-view 3D scene understanding. Our method builds upon recent advances in still-image action recognition and pose estimation, to extract functional and geometric constraints about the scene from people detections. These constraints are then used to improve state-of-the-art single-view 3D scene understanding approaches. The proposed method is validated on a collection of single-viewpoint time-lapse image sequences as well as a dataset of still images of indoor scenes. We demonstrate that observing people performing different actions can significantly improve estimates of scene geometry and 3D layout. The main idea of this work is illustrated in figure 7.
This work is in submission to CVPR 2012.

Figure 7. What can human actions tell us about the 3D structure of the scene? Quite a lot, actually. Consider the two person detections and their estimated pose in (a). They were detected in a time-lapse sequence of one of the three scenes (b-d). Can you guess which one? Most people can easily see that it is (b). Even though this is only a static image, the actions and the pose of the disembodied figures reveal a lot about the geometric structure of the scene. The pose of the left figure reveals a horizontal surface right under its pelvis, which ends abruptly at the knees. The right figure’s pose reveals a ground plane under its feet as well as a likely horizontal surface near the hand location. In both cases we observe a strong physical and functional coupling that exists between people and the 3D geometry of the scene. Our aim in this work is to exploit this coupling.

6.4.6. Joint pose estimation and action recognition in image graphs

Participants: K. Raja [INRIA Rennes], Ivan Laptev, Patrick Perez [Technicolor], L. Osei [INRIA Rennes].

Human analysis in images and video is a hard problem due to the large variation in human pose, clothing, camera view-points, lighting and other factors. While the explicit modeling of this variability is difficult, the huge amount of available person images motivates for the implicit, datadriven approach to human analysis. In this work we aim to explore this approach using the large amount of images spanning a subspace of human appearance. We model this subspace by connecting images into a graph and propagating information through such a graph using a discriminatively trained graphical model. We particularly address the problems of human pose estimation and action recognition and demonstrate how image graphs help solving these problems jointly. We report results on still images with human actions from the KTH dataset.

This work has resulted in a publication [13].

6.5. Creation of the SIERRA project-team

6.5.1. From WILLOW alone to WILLOW and SIERRA

The WILLOW team officially started in the Spring of 2007. From the start, it was clear that machine learning was a key ingredient to new breakthroughs, and our activities have steadily grown in this area. In three short years, WILLOW has grown into a mature group of about 30 people, and it divides its activities between computer vision, machine learning, and the cross-pollination of the two fields, with video as one of the core research areas. We have been very successful, with many publications in all the major international conferences and leading journals in both areas, but we are a large group with very diverse interests, ranging from camera
geometry to statistics, and from image retrieval to bioinformatics applications of structured sparse coding. With the creation of the SIERRA project-team, the core machine learning activities of WILLOW have been transferred to the new group.

The two teams continue collaborating with each other (they remain co-located at the INRIA site in central Paris), but have a sharper focus on their respective computer vision and machine learning activities.

6.5.2. SIERRA

The SIERRA project-team was created by the INRIA on January 1st 2011 and is headed by Francis Bach, who received in 2009 a Jr. ERC grant.