Activity Report Saclay - Île-de-France 2017

Edition: 2018-02-19

List of Inria's Research Teams

Team AMIBIO

Algorithms and Models for Integrative BIOlogy

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions).

RESEARCH CENTER
Saclay - Île-de-France

THEME
Computational Biology

Table of contents

Team AMIBIO

Creation of the Team: 2017 January 01, end of the Team: 2017 December 31

Keywords:

Computer Science and Digital Science:
A3.3.3. - Big data analysis
A3.4.1. - Supervised learning
A3.4.2. - Unsupervised learning
A3.4.5. - Bayesian methods
A5.2. - Data visualization
A5.10.3. - Planning
A6.1.3. - Discrete Modeling (multi-agent, people centered)
A6.1.4. - Multiscale modeling
A6.2.3. - Probabilistic methods
A6.2.4. - Statistical methods
A6.2.6. - Optimization
A6.3.3. - Data processing
A6.3.5. - Uncertainty Quantification
A7.1.3. - Graph algorithms
A8.1. - Discrete mathematics, combinatorics
A8.2. - Optimization
A8.11. - Game Theory
A9.2. - Machine learning

Other Research Topics and Application Domains:
B1.1.1. - Structural biology
B1.1.2. - Molecular biology
B1.1.5. - Genetics
B1.1.6. - Genomics
B1.1.9. - Bioinformatics
B1.1.10. - Mathematical biology
B5.10. - Biotechnology
B9.4.1. - Computer science
B9.4.2. - Mathematics
B9.6. - Reproducibility

1. Personnel

Research Scientists
Yann Ponty [Team leader, CNRS, Researcher]
Mireille Régnier [Ecole polytechnique, Senior Researcher, HDR]

Faculty Members
Jean-Marc Steyaert [Ecole polytechnique, Professor aemeritus, HDR]

Philippe Chassignet [Ecole polytechnique, Associate Professor]

Technical Staff

Pauline Pommeret [Inria engineer, until Jan 2017]

PhD Students

Alice Héliou [Ecole polytechnique, until Aug 2017]
Amélie Héliou [Ecole polytechnique, until Aug 2017]
Juraj Michalik [Inria]
Jorgelindo Moreira Da Veiga [Ecole Polytechnique/CIFRE Soredab]
Pauline Pommeret [Bourse ministérielle, since Oct 2017; Inria engineer, until Jan 2017]
Afaf Saaidi [CNRS]
Antoine Soulé [Ecole Polytechnique & Univ. McGill]
Wei Wang [Univ Paris-Sud, until Jun 2017]

Post-Doctoral Fellow

Christelle Rovetta [Inria, from Sep 2017]

Visiting Scientist

Andrea Tanzer [Ecole polytechnique, Oct 2017]

Administrative Assistant

Evelyne Rayssac [Ecole polytechnique]

2. Overall Objectives

2.1. Overall Objectives

Our project addresses a central question in bioninformatics, namely the molecular levels of organization in the cells. The biological function of macromolecules such as proteins and nucleic acids relies on their dynamic structural nature and their ability to interact with many different partners. Therefore, folding and docking are still major issues in modern structural biology and we currently concentrate our efforts on structure and interactions, and aim at a contribution towards efficient RNA design. With the recent development of computational methods aiming to integrate different levels of information, protein and nucleic acid assemblies studies should provide a better understanding on the molecular processes and machinery occurring in the cell and our research extends to several related issues in comparative genomics.

On the one hand, we study and develop methodological approaches for dealing with macromolecular structures and annotation: the challenge is to develop abstract models that are computationally tractable and biologically relevant. Our approach puts a strong emphasis on the modeling of biological objects using classic formalisms in computer science (languages, trees, graphs...), occasionally decorated and/or weighted to capture features of interest. To that purpose, we rely on the wide array of skills present in our team in the fields of combinatorics, formal languages and discrete mathematics. The resulting models are usually designed to be amenable to a probabilistic analysis, which can be used to assess the relevance of models, or test general hypotheses.

On the other hand, once suitable models are established we apply these computational approaches to several particular problems arising in fundamental molecular biology. One typically aims at designing new specialized algorithms and methods to efficiently compute properties of real biological objects. Tools of choice include exact optimization, relying heavily on dynamic programming, simulations, machine learning and discrete mathematics. As a whole, a common toolkit of computational methods is developed within the group. The trade-off between the biological accuracy of the model and the computational tractability or efficiency is to be addressed in a close partnership with experimental biology groups. One outcome is to provide software or platform elements to predict structural models and functional hypotheses.

Increasingly, our integrative approaches have focused on problems arising in computational structural biology, with a strong focus on Bioinformatics methods focusing on the sequence(s) to structure(s) relationship in RiboNucleic Acids (RNAs). RNAs are versatile biomolecules found in all domains of life, of length ranging from 20-30 nucleotides (nts) in micro RNAs to dozens of thousands nucleotides in certain messenger RNAs or coronaviruses. In most functional families, the structure (or lack thereof) adopted by an RNA is instrumental to its mission(s), and is the object of considerable identifiable pressure throughout evolution. Understanding the structure of RNA and its dynamics leads to testable functional hypotheses. Conversely, a deeper understanding of how function requires the adoption of certain structures leads to models and tools for the rational design of RNAs. AMIBio develops methods and algorithms to predict the dynamics of folding, to make sense of low-dimensional experimental data, perform a rational design of functional RNAs, and detect instances of RNA families within genomic and transcriptomic data.

3. Research Program

3.1. RNA and protein structures

At the secondary structure level, we contributed novel generic techniques applicable to dynamic programming and statistical sampling, and applied them to design novel efficient algorithms for probing the conformational space. Another originality of our approach is that we cover a wide range of scales for RNA structure representation. For each scale (atomic, sequence, secondary and tertiary structure...) cutting-edge algorithmic strategies and accurate and efficient tools have been developed or are under development. This offers a new view on the complexity of RNA structure and function that will certainly provide valuable insights for biological studies.

3.1.1. Discrete representations and complexity

Participants: Yann Ponty, Wei Wang, Antoine Soulé, Juraj Michalik.

*Common activity with J. Waldispühl (McGill) and A. Denise (*LRI*).*

Ever since the seminal work of Zuker and Stiegler, the field of RNA bioinformatics has been characterized by a strong emphasis on the secondary structure. This discrete abstraction of the 3D conformation of RNA has paved the way for a development of quantitative approaches in RNA computational biology, revealing unexpected connections between combinatorics and molecular biology. Using our strong background in enumerative combinatorics, we propose generic and efficient algorithms, both for sampling and counting structures using dynamic programming. These general techniques have been applied to study the sequence-structure relationship [46], the correction of pyrosequencing errors [38], and the efficient detection of multi-stable RNAs (riboswitches) [42], [43].

Increasingly, we develop and study parameterized complexity approaches, based on dynamic programming over a tree decomposition, for several combinatorial problems, including RNA design, structure-sequence alignment (aka threading in the context of proteins). The later problem is at the core of Wei Wang's Phd, successfully defended in Dec 2017. In the context of our probabilistic approaches, often based on random generation, such parameterized algorithms usually follow proofs of hardness for the associated enumeration problems.

3.1.2. RNA design.

Participants: Alice Héliou, Yann Ponty.

Joint project with A. Denise (sc Lri), J. Waldispühl (McGill), D. Barash (Univ. Ben-Gurion), and C. Chauve (Simon Fraser University).

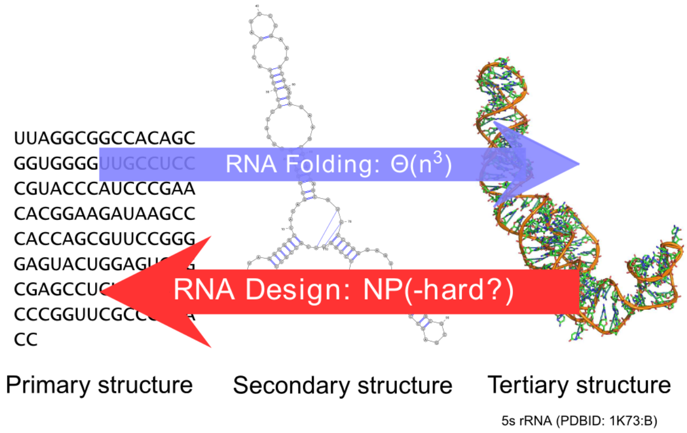


Figure 1. The goal of RNA design, aka RNA inverse folding, is to find a sequence that folds back into a given (secondary) structure.

It is a natural pursuit to build on our understanding of the secondary structure to construct artificial RNAs performing predetermined functions, ultimately targeting therapeutic and synthetic biology applications. Towards this goal, a key element is the design of RNA sequences that fold into a predetermined secondary structure, according to established energy models (inverse-folding problem). Quite surprisingly, and despite two decades of studies of the problem, the computational complexity of the inverse-folding problem is currently unknown.

Within AMIBio, we develop a new methodology, based on weighted random generation [25] and multidimensional Boltzmann sampling, for this problem. Initially lifting the constraint of folding back into the target structure, we explored the random generation of sequences that are compatible with the target, using a probability distribution which favors exponentially sequences of high affinity towards the target. A simple posterior rejection step selects sequences that effectively fold back into the latter, resulting in a *global sampling* pipeline that showed comparable performances to its competitors based on local search [32].

The main advantages of this approach is its linear complexity, and its flexibility in incorporating constraints. Indeed, extensive experiments revealed a drift of existing software towards sequences of high G+C-content, and we showed how to control this distributional bias by using multidimensional Boltzmann sampling [37], [36]. Recently, we are extending this approach to the design of RNAs with multiple structures, developping a Fixed-Parameter Tractable framework that naturally extends to capture negative design goals.

3.1.3. Modeling large macromolecular architectures

Participants: Yann Ponty, Afaf Saaidi, Mireille Régnier, Amélie Héliou.

*Joint projects with A. Denise (*LRI*), D. Barth (Versailles), J. Cohen (Paris-Sud), B. Sargueil (Paris V) and Jérome Waldispühl (McGill).*

The modeling of large RNA 3D structures, that is predicting the three-dimensional structure of a given RNA sequence, relies on two complementary approaches. The approach by homology is used when the structure of a sequence homologous to the sequence of interest has already been resolved experimentally. The main problem then is to calculate an alignment between the known structure and the sequence. The ab initio approach is required when no homologous structure is known for the sequence of interest (or for some parts of it). We contribute methods inspired by both of these directions.

We also develop homology-based approaches for structure modeling, and developed a general setting for the problem of RNA structure-sequence alignment, known to be NP-hard in the presence of complex topological features named pseudoknots (PKs). Our approach is based on tree decomposition of structures and gives rises to a general parameterized algorithm, where the exponential part of the complexity depends on the family of structures [39]. This work unifies and generalizes a number of recent works on specific families, and enables the curation of multiple alignments for RNA families featuring PKs, correcting certain bias introduced by PK-oblivious methods.

3.2. Séquences

Participants: Mireille Régnier, Philippe Chassignet, Yann Ponty, Jean-Marc Steyaert, Alice Héliou, Antoine Soulé.

String searching and pattern matching is a classical area in computer science, enhanced by potential applications to genomic sequences. In CPM/SPIRE community, a focus is given to general string algorithms and associated data structures with their theoretical complexity. Our group specialized in a formalization based on languages, weighted by a probabilistic model. Team members have a common expertise in enumeration and random generation of combinatorial sequences or structures, that are *admissible* according to some given constraints. A special attention is paid to the actual computability of formula or the efficiency of structures design, possibly to be reused in external software.

As a whole, motif detection in genomic sequences is a hot subject in computational biology that allows to address some key questions such as chromosome dynamics or annotation. Among specific motifs involved in molecular interactions, one may cite protein-DNA (cis-regulation), protein-protein (docking), RNA-RNA (miRNA, frameshift, circularisation). This area is being renewed by high throughput data and assembly issues. New constraints, such as energy conditions, or sequencing errors and amplification bias that are technology dependent, must be introduced in the models. A collaboration has beenestablished with LOB, at Ecole Polytechnique, who bought a sequencing machine, through the co-advised thesis of Alice Héliou. An other aim is to combine statistical sampling with a fragment based approach for decomposing structures, such as the cycle decomposition used within F. Major's group [34]. In general, in the future, our methods for sampling and sequence data analysis should be extended to take into account such constraints, that are continuously evolving.

3.2.1. Combinatorial Algorithms and motifs

Participants: Mireille Régnier, Philippe Chassignet, Alice Héliou.

Besides applications [41] of analytic combinatorics to computational biology problems, the team addressed general combinatorial problems on words and fundamental issues on languages and data structures. Motif detection combines an algorithmic search of potential sites and a significance assessment. To assess the significance of an observation usually requires the evaluation of a quantitative criterion such as the P-value. In the recent years, a general scheme of derivation of analytic formula for the P-value under different constraints (k-occurrence, first occurrence, overrepresentation in large sequences,...) has been provided. It relies on a representation of continuous sequences of overlapping words, currently named *clumps* or *clusters* in a graph [35]. Recursive equations to compute p-values may be reduced to a traversal of that graph, leading to a linear algorithm. This improves over the space and time complexity of the generating function approach or previous probabilistic weighted automata.

In [45], it is claimed that half of the genome consists of different types of repeats. One may cite microsatellites, DNA transposons, transposons, long terminal repeats (LTR), long interspersed nuclear elements (LINE), ribosomal DNA, short interspersed nuclear elements (SINE). Therefore, knowledge about the length of repeats is a key issue in several genomic problems, notably assembly or re-sequencing. Preliminary theoretical results are given in [29], and, recently, heuristics have been proposed and implemented [26], [40], [23]. A dual problem is the length of minimal absent words. Minimal absent words are words that do not occur but whose proper factors all occur in the sequence. Their computation is extremly related to finding maximal repeats (repeat that can not be extended on the right nor on the left). The comparison of the sets of minimal absent words provides a fast alternative for measuring approximation in sequence comparison [22], [24].

Recently, it was shown that considering the words which occur in one sequence but do no in another can be used to detect biologically significant events [44]. We have studied the computation of minimal absent words and we have provided new linear implementations [20]. We are now working on a dynamic approach to compute minimal absent words for a sliding window. For a sequence of size n, we expect a complexity of $O(n)$ time and space, independent of the size of the window. This approach could be use to align a sequence on a larger sequence using minimal absent words for comparison.

3.2.2. Random generation

Participants: Yann Ponty, Juraj Michalik, Christelle Rovetta.

Analytical methods may fail when both sequential and structural constraints of sequences are to be modelled or, more generally, when molecular *structures* such as RNA structures have to be handled. The random generation of combinatorial objects is a natural, alternative, framework to assess the significance of observed phenomena. General and efficient techniques have been developed over the last decades to draw objects uniformly at random from an abstract specification. However, in the context of biological sequences and structures, the uniformity assumption becomes unrealistic, and one has to consider non-uniform distributions in order to derive relevant estimates. Typically, context-free grammars can handle certain kinds of long-range interactions such as base pairings in secondary RNA structures.

In 2005, a new paradigm appeared in the *ab initio* secondary structure prediction [27]: instead of formulating the problem as a classic optimization, this new approach uses statistical sampling within the space of solutions. Besides giving better, more robust, results, it allows for a fruitful adaptation of tools and algorithms derived in a purely combinatorial setting.

We also introduced algorithms and data structures for a non-redundant generation of combinatorial objects. In situtations where the search space of a problem can be unambigously explored using dynamic programming, such algorithms generate objects within a postulated distribution, conditioned to avoid previously generated objects. This method can be used to probe objects having lower probabilities, a desirable property in the context of RNA kinetics studies, or could lead to better estimators in context where the exact emission probability of each object can be computed.

3.3. 3D interaction and structure prediction

Participant: Amélie Héliou.

The biological function of macromolecules such as proteins and nucleic acids relies on their dynamic structural nature and their ability to interact with many different partners. This is specially challenging as structure flexibility is key and multi-scale modelling [21], [28] and efficient code are essential [33].

Our project covers various aspects of biological macromolecule structure and interaction modelling and analysis. First protein structure prediction is addressed through combinatorics. The dynamics of these types of structures is also studied using statistical and robotics inspired strategies. Both provide a good starting point to perform 3D interaction modelling, accurate structure and dynamics being essential.

Our group benefits from a good collaboration network, mainly at Stanford University (USA), HKUST (Hong-Kong) and McGill (Canada). The computational expertise in this field of computational structural biology is represented in a few large groups in the world (e.g. Pande lab at Stanford, Baker lab at U.Washington) that have both dry and wet labs. At Inria, our interest for structural biology is shared by the ABS and ORPAILLEUR project-teams. Our activities are however now more centered around protein-nucleic acid interactions, multi-scale analysis, robotics inspired strategies and machine learning than protein-protein interactions, algorithms and geometry. We also shared a common interest for large biomolecules and their dynamics with the NANO-D project team and their adaptative sampling strategy. As a whole, we contribute to the development of geometric and machine learning strategies for macromolecular docking.

Game theory was used by M. Boudard in her PhD thesis, defended in 2015, to predict the 3d structure of RNA. In her PhD thesis, co-advised by J. Cohen (LRI), A. Héliou extended the approach to the prediction of protein structures.

3.3.1. Robotics-inspired structure and dynamics

Participant: Amélie Héliou.

We recently work one a robotics approach to sample the conformational space of macromolecules like RNAs [1]. The robotics approach allows maintaining the secondary structure of the RNA fixed, as an unfolding is very unlikely and energetically demanding. By this approach we also dramatically reduce the number of degrees of freedom in the molecule. The conformational space becomes possible to be sampled. This reduction does not reduce the quality of the sampling.

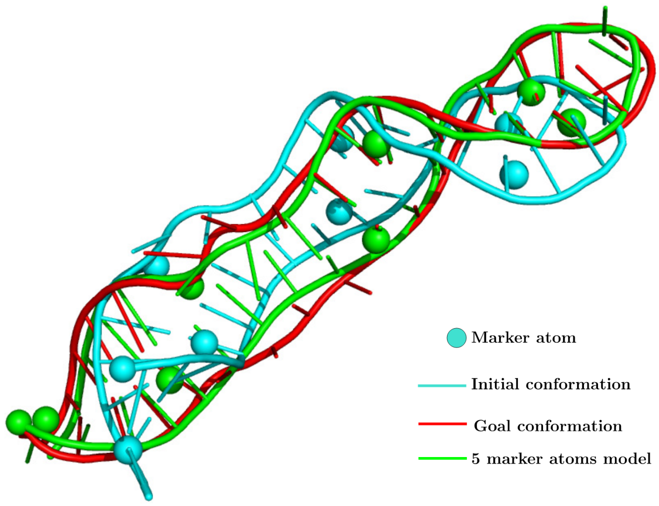


Figure 2. The cyan structure is the initial conformation, the red structure is the goal conformation. The full-atom initial conformation was driven toward the goal conformation using only the position of the goal sphere atoms. The green conformation is the result obtained; spheres perfectly overlap with the goal position and the overall conformation is really close to the goal conformation.

Our current work consists in applying the same approach to a targeted move. The motion is then driven either by the position of a few atoms or the distances between couple of atoms. Theses two aspects are under development and will increase our capacity to integrate experimental data. Our method can drive a RNA conformation toward another conformation of the same RNA given only the position of a few atoms (marker atoms).

For instance double electron-electron resonance (DEER) experimental results are distributions of distances. Probes are attached to the molecules and the distances between to probes is measured and outputted as a distribution. Our method is able to sample an ensemble of all-atom conformations that can explain the distance distribution.

3.3.2. Game theory and molecular folding

Participant: Amélie Héliou.

Building on a previous collaboration with LRI (J. Cohen, expert in algorithmic game theory – GALAC team), we have extended an original approach to structure prediction based on game theory, previously introduced as a proof-of-concept [31]. Our model of the folding model as a game represents an alternative paradigm to thermodynamics and energy minimization, in which individual residues are considers as selfish players, whose strategy (position/orientation) is driven by a desire to increase their local utility function (consistency with observed local conformations). The conformations where residues can no longer gain any utility by changing their local conformations are called Nash equilibria. They can be thought of as local minima of a suitably-defined energy function, and can sometimes be efficiently computed using stochastic strategies.

Our work is first to find an algorithm that can guarantee the convergence to an Nash equilibrium (a state were no player would increase his payoff by playing something different alone) and prove their convergence. At the same time, we are looking for efficient and biologically relevant ways of defining the game settings so that Nash equilibria correspond to folded states. One direction would be to draw a parallel between Nash equilibria and local minima of the kinetic landscape. This line of research also raises questions related to learning Nash equilibria.

4. Application Domains

4.1. Circular RNAs

Participants: Mireille Régnier, Alice Héliou.

Circular RNAs (circRNAs) have been found abundantly in human cells as well as in many other animals. These non-coding RNAs are involved in the regulation of numerous biological processes, and it was recently shown that, as pre-miRNA, they might may actually encode short functional peptides. Our collaborators at Ecole Polytechnique (Biology Dept, LOB) have demonstrated the role of RNA ligase *Pab1020* in RNA circularization. The protein *Pab1020* is a member of the conserved *Rnl3* family of RNA ligases that are predominantly found in hyperthermophiles (archaea, bacteria) and halophiles.

Many computational methods have been proposed to identify and characterize circular RNA from high throughput sequencing data. However, they all suffer from a low specificity, leading to an explosion of false positives. Along with our partners at LOB (Ecole Polytechnique), we develop a robust method for the detection of circRNAs, particularly well-suited to accomodate to analyze sequencing data acquired in extreme environments.

4.2. Analysis of probing data

Participants: Yann Ponty, Mireille Régnier, Afaf Saaidi.

SHAPE probing [47] is an experimental technique in which RNA is exposed to a reagent which, upon reverse-transcription, induces a modification (truncation, mutation) in the DNA. The prevalence of such modifications, which depends on the locally adopted structure(s) (or lack thereof), can be measured for each nucleotide using sequencing techniques, informing regarding the 2D structure. SHAPE probing data can thus be used by structure prediction methods, either to assess their consistency with a proposed structural model, or to restrict the conformation space.

As part of a colloboration with B. Sargueil's lab (Faculté de pharmacie, Paris V) funded by the Fondation pour la Recherche medical, we strive to propose a new paradigm for the analysis data produced using a new experimental technique, called SHAPE analysis (Selective 2'-Hydroxyl Acylation analyzed by Primer Extension). This experimental setup produces an accessibility profile associated with the different positions of an RNA, the *shadow* of an RNA. We currently design new algorithmic strategies to infer the secondary structure of RNA from multiple SHAPE experiments performed by experimentalists at Paris V. Those are obtained on mutants, and will be coupled with a fragment-based 3D modeling strategy developed by our partners at McGill.

5. New Results

5.1. New circular RNAs identified in Pyrococcus abyssi

We contributed a new method for the detection of circRNAs, which we validated on simulated data, and used to analyze the transcriptome of *Pyrococcus abyssi*, an archae living at high depth and temperature [1]. Using this method, which was shown to produce less false positives than previous computational approaches, we analyzed data produced in collaboration with LOB (Ecole Polytechnique), and detected roughly a hundred of novel candidates circular RNAs. Moreover, we provided evidence, on a large scale, that the protein *Pab1020* acts as a ligase, and interacts with some of these circular RNAs, shedding new light on the mechanisms underlying the circularization process.

5.2. Minimal absent words

Minimal absent words are words that do not occur but whose proper factors all occur in the sequence. In a collaboration with King's College, several algorithms, we have designed algorithms to search for minimal absent words in external memory [8], and *in-line*, using a sliding window [13] (parallelization, external memory,...) that outperform previous solutions and achieve near-optimal speed up. This opens new scenarios in the applications of minimal absent words in computational biology, including phylogeny or evolution. For instance, it was shown that there exist three minimal words in Ebola virus genomes which are absent from human genome. As two strings coincide iff they have the same set of minimal absent words, an interesting side result is to solve in optimal time the pattern matching problem using *negative information*.

5.3. Kinematics-inspired algorithms for macromolecular modeling

At a geometric level, RNA is much more flexible than protein, and undergoes smooth transitions between its various conformations. Such transitions are difficult to observe, but can be predicted using algorithms inspired by kinematics and motion-planning. With our partners at Stanford, we designed and implemented such an algorithm within the KGS library [8] to morph between two RNA conformations while keeping distance constraints induced by base pairs and, more importantly, avoiding clashes. In a more preliminary work, we also used similar approaches to automagically fit multi-conformer ligand models into electron density maps [16].

5.4. RNA design

In a paper published in Algorithmica [6], we have shown that our previous results [30] hold for more sophisticated energy models where base-pairs are associated with arbitrary energy contributions. This result, which required a complete overhaul of our previous proofs (e.g. using arguments based on graph coloring), allows us to foresee an extension of (at least some of) our results to state-of-the-art models, such as the Turner energy model.

In collaboration with Danny Barash's group at Ben-Gurion university (Israel), we contributed a review of existing tools and techniques for RNA design, which was published in *Briefings in Bioinformatics* series [3].

Finally, in a paper [14] recently accepted for a presentation at the prestigious RECOMB'18 conference, we revisited the problem of generating at random an RNA sequence which is simultaneously compatible with a set of target secondary structures. This problem was previously addressed by our collaborators at the TBI Vienna/Univ. Leipzig, using an exponential-time algorithm. We established the #*P*-hardness of the problem, and its inapproximibility in general. However, the problem is still amenable to an efficient parametrization, and we proposes an FTP algorithm named `RNARedPrint` based on the tree decomposition for the random generation, to which we adapted a multidimensionnal Boltzmann sampling technique in order to gain (probabilistic) control over secondary features such as the *GC%*, the relative free-energy of the various structures...

5.5. Game theory and macromolecular modeling

Initially based on a very coarse representation of RNA, we refined our model of RNA folding as a game, using on-lattice coordinates and statistical potentials for the utility function. The resulting algorithm was implemented in the subsequent version of the `GARN` [2] software.

The final year of Amélie Héliou's PhD led to theoretical developments in game theory, mainly obtained in collaboration with J. Cohen (LRI, Univ. Paris-Sud). First, the quasi-exponential convergence, under reasonable assumptions, of the HEDGE algorithm was demonstrated in collaboration with the POLARIS team in Grenoble [11]. Moreover, in a paper accepted at NIPS'17 [12], we addressed the learning of Nash equilibria. In this context, we established the convergence with high probability of no-regret learning in the bandit and semi-bandit settings.

5.6. RNA kinetics using non-redundant sampling

RNA kinetics is arguably the next frontier in RNA 2D bioinformatics. In particular, computational methods for studying the kinetics of RNA beyond 150nts are hindered by the combinatorial explosion of the conformation space. In an effort to circumvent such an effect, we have proposed a sampling approach that explicitly target local minima of the energy function. Our sampling algorithm, jointly proposed with H. Touzet (Bonsai, Inria Lille & CrisTaL, Univ. Lille I) and accepted for a presentation at the ISMB/ECCB'17 conference in Prague [15], uses non-redundant sampling principles to avoid an excessive concentration of samples within low local minima.

5.7. New insight from SHAPE probing data

Existing computational methods for structure prediction are typically hindered by their assumption of a single structure, and their assumption of orthogonal signals stemming from different reagents. To overcome these limitations, we contributed an integrative approach combining stochastic sampling and structural clustering [17] (journal version pending). In collaboration with ENS Lyon/Univ. Lyon I and Univ. Paris-Descartes, we used this method to model the structure of the HIV-1 gag open-reading frame [4].

We also addressed the problem of binning sets of NGS reads arising from the simultaneous probing, using the SHAPEmap protocol, of variants produced by a error-prone PCR. We proposed a variant of the Expectation-Maximization algorithm [10] to jointly infer maximum-likelihood origins for reads and mutational profiles for each variant.

6. Partnerships and Cooperations

6.1. National Initiatives

6.1.1. FRM

AMIBio is in charge of Bioinformatics developments in this project on structural prediction from RNA probing data (SHAPE). It involves Biochemists at Université Paris Descartes (France, PI B. Sargueil) and is funded by a "Fondation pour la Recherche Medicale" grant. It also involves partners in Paris-Sud (France) and McGill University (Canada).

Fondation pour la Recherche Medicale – *Analyse Bio-informatique pour la recherche en Biologie* program

- Approche comparatives haut-débit pour la modelisation de l'architecture 3D des ARN à partir de données experimentales
- 2015–2018
- Yann Ponty, A. Denise, M. Regnier, A. Saaidi (PhD funded by FRM)
- B. Sargueil (Paris V – Experimental partner), J. Waldispuhl (Univ. McGill)

6.2. European Initiatives

6.2.1. Collaborations in European Programs, Except FP7 & H2020

Yann Ponty is the French PI for the French/Austrian RNALANDS project, jointly funded by the French ANR and the Austrian FWF, in partnership with the Theoretical Biochemistry Institute (University of Vienna, Austria), LRI (Univ. Paris-Sud) and EPI BONSAI (Inria Lille-Nord Europe).

French/Austrian International Program

RNALANDS (ANR-14-CE34-0011)

Fast and efficient sampling of structures in RNA folding landscapes

01/10/2014–30/09/2018

Coordinated by AMIB (Inria Saclay) and TBI Vienna (University of Vienna)

EPI BONSAI/INRIA Lille - Nord Europe, Vienna University (Austria), LRI, Université Paris-Sud (France)

The main goal of the RNALands project is to provide efficient tools for studying the kinetics of RiboNucleic Acids, based on efficient sampling strategies.

6.3. International Initiatives

6.3.1. Inria Associate Teams Not Involved in an Inria International Labs

6.3.1.1. ALARNA

Title: Associated Laboratory for the Analysis of RiboNucleic Acids

International Partner (Institution - Laboratory - Researcher):

McGill University (Canada) - REUSSI Program - Jerome Waldispuhl

Start year: 2017

See also: https://team.inria.fr/alarna/

RiboNucleic Acids (RNAs) are ubiquitous biomolecules whose structure, adopted as the outcome of a complex folding process, often plays a crucial part in cellular processes. The ALARNA Associate Team (Laboratory for the Analysis of RiboNucleic Acids), which consist of the AMIBio project-team (Inria Saclay/Ecole Polytechnique, France) and the CSB (Computer Science and Biology) group at university McGill (Montreal, Canada), addresses key questions in RNA bioinformatics. More specifically, it dedicates much of its effort to the production and interpretation of chemical probing data generated by SHAPE, an experimental technology which allows to accurately predict, in a high-throughput, one or several secondary structure(s) adopted by an RNA. To that end, the teams contribute their unique combinations of expertise, ranging from combinatorial optimization to sequence algorithmics through structural bioinformatics.

6.3.1.2. Informal International Partners

AMIBio enjoys regular interactions with the following institutions:

- TBI, University of Vienna (Austria). Within the RNALands project funded by the Austrian FWF and the french ANR, we frequently interact with our partners at the TBI, on projects associated with the kinetics of RNAs. Over the course of 2017, we have visited our partners twice, once in Vienna and once in Bled (Slovenia) over the course of the 2017 Winter retreat of the TBI. Additionally, Andrea Tanzer has visited AMIBio for a month in Oct 2018, funded by a visiting scholar program of Ecole Polytechnique;
- Simon Fraser University (Vancouver, Canada). The Mathematics department at SFU has ongoing projects on RNA design, comparative genomics and RNA structure comparison with our team. M. Mishna (SFU) has visited Inria Saclay in January 2017 to push an ongoing collaboration on 2D walks;

- McGill University (Montréal, Canada). Following our productive collaboration with J. Waldispühl (Computer Science Dept, McGill), and the recent defense of V. Reinharz's PhD, whose thesis was co-supervised by AMIBio members, we have increased our interactions on SHAPE data analysis through the ALARNA associate team;
- King's college (London, UK). Our collaboration with L. Mouchard (AMIBio associate) and S. Pissis on string processing and data structures was at the core of Alice Héliou's PhD, defended in July 2017.

6.3.2. Participation in Other International Programs

Title: PHC GRO-algo – Combination of time-course GRO-seq assay, algorithmics and software development for measuring genome-wide transcription elongation rates

International Partner (Institution - Laboratory - Researcher):

Wuhan University (China), College of Life Science – Pr Yu Zhou

Start year: 2017

Participant in a French-Chinese Hubert Curien Partnerships (PHC), supported by CampusFrance and funding bilateral exploratory research exchanges in Bioinformatics. The program involves research scientists from Wuhan University, Ecole Polytechnique and Univ. Paris-Sud.

Title: Computational methods and databases to identify small RNA-binding molecules regulating gene expression

International Partner (Institution - Laboratory - Researcher):

University McGill (Canada), Computer Science & Biochemistry – J. Waldispühl, N. Moitessier; Univ. Strasbourg, IBMC - E. Westhof.

Start year: 2017

The project, headed by N. Moitessier and J. Waldispühl (McGill University, Canada) strives to develop tools to derive a mechanical understanding of riboswitches at the 2D and 3D levels, including chemoinformatics aspects.

6.4. International Research Visitors

6.4.1. Visits of International Scientists

Andrea Tanzer

Date: Oct 2017 - Nov 2017

Institution: TBI Vienna, Austria

Mathieu Blanchette

Date: June 2017

Institution: Univ. McGill, Canada

6.4.1.1. Internships

Paul Arijit

Institution: IISc Bangalore (India)

Supervisor: Mireille Régnier

Chinmay Singhal

Date: May 2017 - July 2017

Institution: IIT Guwahati, India (India)

Supervisor: Yann Ponty

7. Dissemination

7.1. Promoting Scientific Activities

7.1.1. Scientific Events Organisation

7.1.1.1. General Chair, Scientific Chair

- RECOMB'18: Y. Ponty, M. Regnier (co-chair)

7.1.2. Scientific Events Selection

7.1.2.1. Member of the Conference Program Committees

- RECOMB'17: Y. Ponty, M. Regnier
- ISMB/ECCB'17: Y. Ponty
- ACM/BCB'17: A. Héliou
- MCCMB'17: M. Régnier
- BiCoB'17: Y. Ponty

7.1.2.2. Reviewer

- CPM'17: Y. Ponty

7.1.3. Journal

7.1.3.1. Member of the Editorial Boards

M. Régnier is an editor of PeerJ Computer Science.

7.1.3.2. Reviewer - Reviewing Activities

M. Régnier and Y. Ponty reviewed manuscripts for a large selection of journals in Mathematics, Computer Science and Bioinformatics: Discrete Mathematics and Theoretical Computer Science, Theoretical Computer Science, Bioinformatics, BMC Bioinformatics, Journal of Mathematical Biology, IEEE/ACM Transactions on Computational Biology and Bioinformatics, Journal of Discrete Algorithms, Algorithms for Molecular Biology, PLOS One, Journal of Theoretical Biology, RNA, Nucleic Acids Research...

7.1.4. Leadership within the Scientific Community

Y. Ponty is *animateur* of the *Macromolecular structure and interactions* axis of the CNRS GDR BIM (BioInformatique Moléculaire). With F. Cazals (ABS, Inria Sophia-Antipolis), he co-created in Oct. 2017 and currently heads the MASIM (Méthodes Algorithmiques pour les Structures et Interactions des Macromolécules) workgroup of BIM;

M. Régnier is a member of DIGITEO progam Committee and SDV working group in Saclay area.

7.1.5. Scientific Expertise

Y. Ponty reviewed grants for the OPUS program of the NCN agency (Poland), the Bioinformatics and Theoretical Biology programm of the DFG (Germany). He acted as an external reviewer and external assessor for an assistant professor position at the Faculty of Science of the University of South Denmark. He also refereed for a postdoc call by the French Charmmmat LabEx of Univ. Paris-Saclay.

7.1.6. Research Administration

M. Regnier is the current head of the Computer Science department (LIX) of Ecole Polytechnique.

7.2. Teaching - Supervision - Juries

7.2.1. Teaching

7.2.1.1. Initial training in engineering at Ecole Polytechnique.

At the undergraduate level, AMIBio is essentially involved in Computer Science courses at Ecole Polytechnique, mainly in programming languages and Big Data processing (high-performance computing and machine learning). Notably, we are involved in the *parcours d'Approfondissement en Bioinformatique* at École Polytechnique.

7.2.1.2. Graduate-level programs and courses.

Our project is also very much involved in the AMI2B (*Analyse, Modélisation et Ingénierie de l'Information Biologique et Médicale*, formerly named BIBS – *Bioinformatique et Biostatistique*) Master program at Université Paris-Sud/École Polytechnique. Most AMIBio permanent members teach recurrent courses in AMI2B, and M. Regnier is in charge of the program at the M1 and M2 levels for Ecole Polytechnique.

Beyond the *plateau de Saclay*, AMIBio members participate in the BIM master program at UPMC, and regularly deliver PhD-level lectures in Summer/Winter schools.

Bachelor (Licence & Ecole Ingénieur)

- P. Chassignet, INF*, 100h, L2/L3, Ecole Polytechnique, France
- Al. Héliou, INF*, 40h, L3, Ecole Polytechnique, France

Master

- P. Chassignet, INF*/Modal Bioinfo, 90h, M1/M2, Ecole Polytechnique, France
- Al. Héliou, INF*, 7h, M1 AMI2B, Univ. Paris-Saclay, France
- Am. Héliou, INF*, 44h, M1 AMI2B, Univ. Paris-Saclay, France
- Y. Ponty, Bioinfo ARN, 12h AMI2B, M2, Univ. Paris-Saclay, France
- Y. Ponty, Bioinfo ARN, 12h BIM, M2, UPMC, France
- M. Regnier, Algo/Combinatoire, 12h AMI2B, M2, Univ. Paris-Saclay, France
- J.-M. Steyaert, Algo/Combinatoire, 12h AMI2B, M2, Univ. Paris-Saclay, France
- J.-M. Steyaert, Bioinformatics INF/BIO 588, 20h Majeure BioInfo, M1, Ecole Polytechnique,France

Doctorat

- Y. Ponty, Design ARN, 3h BIM, PhD, Univ. Tehran, Iran

7.2.2. Supervision

PhD :

Wei Wang, *Practical structure-sequence alignment of pseudoknotted RNAs*, Univ. Paris-Saclay, December 2017, Y. Ponty, A. Denise

Alice Héliou, *Analyse des séquences génomiques : Identification des ARNs circulaires et calcul de l'information négative*, Univ. Paris-Saclay, July 2017, M. Regnier

Amélie Héliou, *Théorie des jeux et échantillonnage de conformations pour l'amarrage macromoléculaire multi-corps et multi-échelle*, Univ. Paris-Saclay, August 2017, J. Cohen

PhD in progress

Afaf Saaidi, *Differential analysis of RNA SHAPE probing data*, Ecole Polytechnique, Encadrants: Yann Ponty and Mireille Régnier.

Antoine Soulé, *Evolutionary study of RNA-RNA interactions in yeast* , Ecole Polytechnique, Encadrants: Jean-Marc Steyaert and J. Waldispuhl (U. McGill, Canada);

Jorgelindo Moreira da Veiga, *Caractérisation dynamique et optimisation des flux métaboliques* , Ecole Polytechnique, Encadrants: L. Schwartz (AP-HP)Sabine Peres (U. Paris-Sud)

Juraj Michalik, *Non-redundant sampling for the study of RNA kinetics*, Inria Saclay, Encadrants: Y. Ponty and H. Touzet (Cristal, Univ. Lille I);

Pauline Pommeret, *Étude de l'impact phénotypique de variants génétiques altérant la structure secondaire des ARNs chez le bovin : Méthodes algorithmiques et validations expérimentales*, INRA Jouy & Ecole Polytechnique, Encadrants: Y. Ponty and D. Rocha (INRA Jouy);

Ha Thi Ngoc Nguyen, *New computational strategies for analyzing the diversity of transcriptome sequence and structure, and their relationship to disease*, I2BC (Univ. Paris-Sud) & Ecole Polytechnique, Encadrants: Y. Ponty and D. Gautheret (Univ. Paris-Sud);

7.2.3. Juries

Y. Ponty participated in a hiring committee for an associate professor position in Discrete Mathematics at the University of Southern Denmark;

Y. Ponty acted as an external opponent for the PhD defence of X. Pan at the University of Copenhagen;

7.3. Popularization

May 2017 Participation to the CURIOSITas art & science festival organized by univ. Paris-Saclay on the CNRS campus (Gif-sur-Yvette, France);

March 2017 Outreach presentation by AMIBio on RNA folding and combinatorial design for a class of high school students (Montmorency, France);

8. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] H. F. BECKER, A. HÉLIOU, K. DJAOUT, R. LESTINI, M. REGNIER, H. MYLLYKALLIO.*High-Throughput Sequencing Reveals Circular Substrates for an Archaeal RNA ligase*, in "RNA Biology", March 2017, https://hal.archives-ouvertes.fr/hal-01491132.

[2] M. BOUDARD, D. BARTH, J. BERNAUER, A. DENISE, J. COHEN.*GARN2: coarse-grained prediction of 3D structure of large RNA molecules by regret minimization*, in "Bioinformatics", 2017, vol. 16, p. 2479-2486 [*DOI :* 10.1093/BIOINFORMATICS/BTX175], https://hal.archives-ouvertes.fr/hal-01589347.

[3] A. CHURKIN, M. D. RETWITZER, V. REINHARZ, Y. PONTY, J. WALDISPÜHL, D. BARASH.*Design of RNAs: comparing programs for inverse RNA folding*, in "Briefings in Bioinformatics", January 2017 [*DOI :* 10.1093/BIB/BBW120], https://hal.inria.fr/hal-01392958.

[4] J. DEFORGES, S. DE BREYNE, M. AMEUR, N. ULRYCK, N. CHAMOND, A. SAAIDI, Y. PONTY, T. OHLMANN, B. SARGUEIL.*Two ribosome recruitment sites direct multiple translation events within HIV1 Gag open reading frame*, in "Nucleic Acids Research", July 2017, vol. 45, n^o 12, p. 7382–7400 [*DOI :* 10.1093/NAR/GKX303], https://hal.archives-ouvertes.fr/hal-01505282.

[5] W. DUCHEMIN, Y. ANSELMETTI, M. PATTERSON, Y. PONTY, S. BÉRARD, C. CHAUVE, C. SCORNAVACCA, V. DAUBIN, E. TANNIER.*DeCoSTAR: Reconstructing the ancestral organization of genes or genomes using reconciled phylogenies*, in "Genome Biology and Evolution", 2017, vol. 9, n^o 5, p. 1312-1319, https://hal.inria.fr/hal-01503766.

[6] J. HALEŠ, A. HÉLIOU, J. MAŇUCH, Y. PONTY, L. STACHO.*Combinatorial RNA Design: Designability and Structure-Approximating Algorithm in Watson-Crick and Nussinov-Jacobson Energy Models*, in "Algorithmica", November 2017, vol. 79, n^o 3, p. 835–856, https://arxiv.org/abs/1603.03577 [*DOI :* 10.1007/s00453-016-0196-X], https://hal.inria.fr/hal-01285499.

[7] A. HÉLIOU, D. BUDDAY, R. FONSECA, H. VAN DEN BEDEM.*Fast, clash-free RNA conformational morphing using molecular junctions*, in "Bioinformatics", July 2017, vol. 33, n^{o} 14, p. 2114 - 2122 [*DOI :* 10.1093/BIOINFORMATICS/BTX127], https://hal.archives-ouvertes.fr/hal-01569620.

[8] A. HÉLIOU, S. P. PISSIS, S. J. PUGLISI.*emMAW: Computing Minimal Absent Words in External Memory*, in "Bioinformatics", 2017 [*DOI :* 10.1093/BIOINFORMATICS/BTX209], https://hal.archives-ouvertes.fr/hal-01569271.

[9] B. LÖWES, C. CHAUVE, Y. PONTY, R. GIEGERICH.*The BRaliBase dent - a tale of benchmark design and interpretation*, in "Briefings in Bioinformatics", March 2017, vol. 18, n^{o} 2, p. 306–311 [*DOI :* 10.1093/BIB/BBW022], https://hal.inria.fr/hal-01273406.

Invited Conferences

[10] A. SAAIDI, Y. PONTY, M. BLANCHETTE, M. REGNIER, B. SARGUEIL.*An EM algorithm for mapping short reads in multiple RNA structure probing experiments*, in "Matbio2017", London, United Kingdom, King's College London , September 2017, https://hal.inria.fr/hal-01590528.

International Conferences with Proceedings

[11] J. COHEN, A. HÉLIOU, P. MERTIKOPOULOS.*Hedging under uncertainty: regret minimization meets exponentially fast convergence*, in "Symposium on Algorithmic Game Theory (SAGT) 2017", L'Aquila, Italy, Proceedings of the 10th International Symposium on Algorithmic Game Theory, September 2017, https://arxiv.org/abs/1607.08863 [*DOI :* 10.1007/978-3-319-66700-3_20], https://hal.archives-ouvertes.fr/hal-01382290.

[12] J. COHEN, A. HÉLIOU, P. MERTIKOPOULOS.*Learning with bandit feedback in potential games*, in "NIPS '17: Proceedings of the 31st International Conference on Neural Information Processing Systems", Long Beach, CA, United States, December 2017, https://hal.archives-ouvertes.fr/hal-01643352.

[13] M. CROCHEMORE, A. HÉLIOU, G. KUCHEROV, L. MOUCHARD, S. P. PISSIS, Y. RAMUSAT.*Minimal absent words in a sliding window & applications to on-line pattern matching*, in "FCT 2017", Bordeaux, France, Lecture Notes in Computer Science, Springer, September 2017, https://hal.archives-ouvertes.fr/hal-01569264.

[14] S. HAMMER, Y. PONTY, W. WANG, S. WILL.*Fixed-Parameter Tractable Sampling for RNA Design with Multiple Target Structures*, in "RECOMB 2018 – 22nd Annual International Conference on Research in Computational Molecular Biology", Paris, France, April 2018, https://hal.inria.fr/hal-01631277.

[15] J. MICHÁLIK, H. TOUZET, Y. PONTY.*Efficient approximations of RNA kinetics landscape using non-redundant sampling*, in "ISMB/ECCB - 25th Annual international conference on Intelligent Systems for Molecular Biology/16th European Conference on Computational Biology - 2017", Prague, Czech Republic, July 2017, vol. 33, n^{o} 14, p. i283 - i292 [*DOI :* 10.1093/BIOINFORMATICS/BTX269], https://hal.inria.fr/hal-01500115.

[16] G. V. ZUNDERT, D. KEEDY, P. SURESH, A. HÉLIOU, K. BORRELLI, T. DAY, J. FRASER, H. VAN DEN BEDEM.*Objectively and automatically building multi-conformer ligand models in electron densities*, in "Conformational ensembles from experimental data and computer simulations", Berlin, Germany, August 2017, https://hal.inria.fr/hal-01569829.

Conferences without Proceedings

[17] A. SAAIDI, Y. PONTY, B. SARGUEIL.*An integrative approach for predicting the RNA secondary structure for the HIV–1 Gag UTR using probing data*, in "JOBIM 2017 - Journées Ouvertes en Biologie, Informatique et Mathématiques", Lille, France, July 2017, 1, https://hal.archives-ouvertes.fr/hal-01534587.

Other Publications

[18] C. CHAUVE, J. COURTIEL, Y. PONTY.*Counting, generating, analyzing and sampling tree alignments*, 2017, Submitted to IJFCS, https://hal.inria.fr/hal-01500116.

[19] D. SURUJON, Y. PONTY, P. CLOTE.*Small-world networks and RNA secondary structures*, January 2017, working paper or preprint, https://hal.inria.fr/hal-01424452.

References in notes

[20] C. BARTON, A. HELIOU, L. MOUCHARD, S. PISSIS.*Linear-time computation of minimal absent words using suffix array*, in "BMC Bioinformatics", 2014, vol. 15, 11 [*DOI :* 10.1186/S12859-014-0388-9], https://hal.inria.fr/hal-01110274.

[21] J. BERNAUER, S. C. FLORES, X. HUANG, S. SHIN, R. ZHOU.*Multi-Scale Modelling of Biosystems: from Molecular to Mesocale - Session Introduction*, in "Pacific Symposium on Biocomputing", 2011, p. 177-80 [*DOI :* 10.1142/9789814335058_0019], http://hal.inria.fr/inria-00542791.

[22] S. CHAIRUNGSEE, M. CROCHEMORE.*Using minimal absent words to build phylogeny*, in "Theoretical Computer Science", 2012, vol. 450, n° 0, p. 109-116.

[23] R. CHIKHI, P. MEDVEDEV.*Informed and automated k-mer size selection for genome assembly*, in "Bioinformatics", Jan 2014, vol. 30, n° 1, p. 31–37, http://dx.doi.org/10.1093/bioinformatics/btt310.

[24] M. CROCHEMORE, G. FICI, R. MERCAS, S. PISSIS.*Linear-Time Sequence Comparison Using Minimal Absent Words*, in "LATIN 2016: Theoretical Informatics - 12th Latin American Symposium", Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2016, http://arxiv.org/abs/1506.04917.

[25] A. DENISE, Y. PONTY, M. TERMIER.*Controlled non uniform random generation of decomposable structures*, in "Theoretical Computer Science", 2010, vol. 411, n° 40-42, p. 3527-3552 [*DOI :* 10.1016/J.TCS.2010.05.010], http://hal.inria.fr/hal-00483581.

[26] H. DEVILLERS, S. SCHBATH.*Separating significant matches from spurious matches in DNA sequences*, in "Journal of Computational Biology", 2012, vol. 19, n° 1, p. 1–12, http://dx.doi.org/10.1089/cmb.2011.0070.

[27] Y. DING, C. CHAN, C. LAWRENCE.*RNA secondary structure prediction by centroids in a Boltzmann weighted ensemble*, in "RNA", 2005, vol. 11, p. 1157–1166.

[28] S. C. FLORES, J. BERNAUER, S. SHIN, R. ZHOU, X. HUANG.*Multiscale modeling of macromolecular biosystems*, in "Briefings in Bioinformatics", July 2012, vol. 13, n° 4, p. 395-405 [*DOI :* 10.1093/BIB/BBR077], http://hal.inria.fr/hal-00684530.

[29] Z. GU, H. WANG, A. NEKRUTENKO, W. H. LI.*Densities, length proportions, and other distributional features of repetitive sequences in the human genome estimated from 430 megabases of genomic sequence*, in "Gene", Dec 2000, vol. 259, n° 1-2, p. 81–88.

[30] J. HALEŠ, J. MAŇUCH, Y. PONTY, L. STACHO.*Combinatorial RNA Design: Designability and Structure-Approximating Algorithm*, in "Annual Symposium on Combinatorial Pattern Matching", Springer, 2015, p. 231–246.

[31] A. LAMIABLE, F. QUESSETTE, S. VIAL, D. BARTH, A. DENISE.*An algorithmic game-theory approach for coarse-grain prediction of RNA 3D structure*, in "IEEE/ACM Transactions on Computational Biology and Bioinformatics", 2013, vol. 10, n^{o} 1, p. 193-199, http://hal.inria.fr/hal-00756340.

[32] A. LEVIN, M. LIS, Y. PONTY, C. W. O'DONNELL, S. DEVADAS, B. BERGER, J. WALDISPÜHL.*A global sampling approach to designing and reengineering RNA secondary structures*, in "Nucleic Acids Research", November 2012, vol. 40, n^{o} 20, p. 10041-52 [*DOI :* 10.1093/NAR/GKS768], http://hal.inria.fr/hal-00733924.

[33] S. LORIOT, F. CAZALS, J. BERNAUER.*ESBTL: efficient PDB parser and data structure for the structural and geometric analysis of biological macromolecules*, in "Bioinformatics", April 2010, vol. 26, n^{o} 8, p. 1127-8 [*DOI :* 10.1093/BIOINFORMATICS/BTQ083], http://hal.inria.fr/inria-00536404.

[34] M. PARISIEN, F. MAJOR.*The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data*, in "Nature", 2008, vol. 452, n^{o} 7183, p. 51–55.

[35] M. REGNIER, E. FURLETOVA, M. ROYTBERG, V. YAKOVLEV.*Pattern occurrences Pvalues, Hidden Markov Models and Overlap Graphs*, 2013, submitted, http://hal.inria.fr/hal-00858701.

[36] V. REINHARZ, Y. PONTY, J. WALDISPÜHL.*A weighted sampling algorithm for the design of RNA sequences with targeted secondary structure and nucleotide distribution*, in "Bioinformatics", July 2013, vol. 29, n^{o} 13, i308 [*DOI :* 10.1093/BIOINFORMATICS/BTT217], http://hal.inria.fr/hal-00840260.

[37] V. REINHARZ, Y. PONTY, J. WALDISPÜHL.*A weighted sampling algorithm for the design of RNA sequences with targeted secondary structure and nucleotides distribution*, in "ISMB/ECCB - 21st Annual international conference on Intelligent Systems for Molecular Biology/12th European Conference on Computational Biology - 2013", Berlin, Allemagne, 2013, http://hal.inria.fr/hal-00811607.

[38] V. REINHARZ, Y. PONTY, J. WALDISPÜHL.*Using Structural and Evolutionary Information to Detect and Correct Pyrosequencing Errors in Noncoding RNAs*, in "Journal of Computational Biology", November 2013, vol. 20, n^{o} 11, p. 905-19, Extended version of RECOMB'13 [*DOI :* 10.1089/CMB.2013.0085], http://hal.inria.fr/hal-00828062.

[39] P. RINAUDO, Y. PONTY, D. BARTH, A. DENISE.*Tree decomposition and parameterized algorithms for RNA structure-sequence alignment including tertiary interactions and pseudoknots*, in "WABI - 12th Workshop on Algorithms in Bioinformatics - 2012", Ljubljana, Slovénie, B. RAPHAEL, J. TANG (editors), University of Ljubljana, 2012, http://hal.inria.fr/hal-00708580.

[40] G. RIZK, D. LAVENIER, R. CHIKHI.*DSK: k-mer counting with very low memory usage*, in "Bioinformatics", Mar 2013, vol. 29, n^{o} 5, p. 652–653 [*DOI :* 10.1093/BIOINFORMATICS/BTT020], http://bioinformatics.oxfordjournals.org/content/early/2013/02/01/bioinformatics.btt020.full.

[41] C. SAULE, M. REGNIER, J.-M. STEYAERT, A. DENISE.*Counting RNA pseudoknotted structures*, in "Journal of Computational Biology", October 2011, vol. 18, n^{o} 10, p. 1339-1351 [*DOI :* 10.1089/CMB.2010.0086], http://hal.inria.fr/inria-00537117.

[42] E. SENTER, S. SHEIKH, I. DOTU, Y. PONTY, P. CLOTE.*Using the Fast Fourier Transform to Accelerate the Computational Search for RNA Conformational Switches*, in "PLoS ONE", December 2012, vol. 7, n^{o} 12 [*DOI :* 10.1371/JOURNAL.PONE.0050506], http://hal.inria.fr/hal-00769740.

[43] E. SENTER, S. SHEIKH, I. DOTU, Y. PONTY, P. CLOTE.*Using the Fast Fourier Transform to accelerate the computational search for RNA conformational switches (extended abstract)*, in "RECOMB - 17th Annual International Conference on Research in Computational Molecular Biology - 2013", Beijing, Chine, 2013, http://hal.inria.fr/hal-00766780.

[44] R. M. SILVA, D. PRATAS, L. CASTRO, A. J. PINHO, P. J. S. G. FERREIRA.*Three minimal sequences found in Ebola virus genomes and absent from human DNA*, in "Bioinformatics", 2015 [*DOI :* 10.1093/BIOINFORMATICS/BTV189].

[45] T. J. TREANGEN, S. L. SALZBERG.*Repetitive DNA and next-generation sequencing: computational challenges and solutions*, in "Nat Rev Genet", Jan 2012, vol. 13, n^{o} 1, p. 36–46, http://dx.doi.org/10.1038/nrg3117.

[46] J. WALDISPÜHL, Y. PONTY.*An unbiased adaptive sampling algorithm for the exploration of RNA mutational landscapes under evolutionary pressure*, in "Journal of Computational Biology", November 2011, vol. 18, n^{o} 11, p. 1465-79 [*DOI :* 10.1089/CMB.2011.0181], http://hal.inria.fr/hal-00681928.

[47] K. A. WILKINSON, E. J. MERINO, K. M. WEEKS.*Selective 2 [prime]-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution*, in "Nature protocols", 2006, vol. 1, n^{o} 3, p. 1610–1616.

Project-Team AVIZ

Analysis and Visualization

RESEARCH CENTER
Saclay - Île-de-France

THEME
Interaction and visualization

Table of contents

Project-Team AVIZ

Creation of the Team: 2007 February 08, updated into Project-Team: 2008 January 01

Keywords:

Computer Science and Digital Science:

A1.3. - Distributed Systems
A1.3.1. - Blockchain
A3.1.4. - Uncertain data
A3.1.7. - Open data
A3.1.8. - Big data (production, storage, transfer)
A3.3. - Data and knowledge analysis
A3.3.1. - On-line analytical processing
A3.3.3. - Big data analysis
A3.5.1. - Analysis of large graphs
A5.1. - Human-Computer Interaction
A5.2. - Data visualization
A6.3.3. - Data processing

Other Research Topics and Application Domains:

B1. - Life sciences
B1.1. - Biology
B1.2. - Neuroscience and cognitive science
B9.4.5. - Data science
B9.5. - Humanities
B9.5.1. - Psychology
B9.5.3. - Economy, Finance
B9.5.6. - Archeology, History
B9.5.10. - Digital humanities

1. Personnel

Research Scientists

Jean-Daniel Fekete [Team leader, Inria, Senior Researcher, HDR]
Pierre Dragicevic [Inria, Researcher]
Petra Isenberg [Inria, Researcher]
Tobias Isenberg [Inria, Senior Researcher, HDR]

External Collaborators

Evelyne Lutton [INRA]
Frédéric Vernier [Université Paris-Sud]

Technical Staff

Romain Di Vozzo [Inria]

PhD Students

Lonni Besançon [Univ Paris-Sud]
Evanthia Dimara [Inria, until Nov 2017]

Sarkis Halladjian [Inria, from Sep 2017]
Xiyao Wang [Inria, from Nov 2017]

Post-Doctoral Fellows

Christoph Kinkeldey [Inria]
Tanja Blascheck [Inria, from Nov 2017]

Visiting Scientists

Paolo Buono [University of Bari Aldo Moro, until Feb 2017]
Jaemin Jo [Seoul National University, until Feb 2017]
Paola Tatiana Llerena Valdivia [USP - Universidade de São Paulo, until Jul 2017]
Nicola Pezzotti [Delft University, from Apr 2017 until Jun 2017]
Catherine Plaisant [University of Maryland, from May 2017 until Jul 2017]
Vanessa Serrano Molinero [IQS School of Management. Universitat Ramon Llull, from Jun 2017 until Sep 2017]

Administrative Assistant

Katia Evrat [Inria]

2. Overall Objectives

2.1. Objectives

Aviz (Analysis and VIsualiZation) is a multidisciplinary project that seeks to improve visual exploration and analysis of large, complex datasets by tightly integrating analysis methods with interactive visualization.

Our work has the potential to affect practically all human activities for and during which data is collected and managed and subsequently needs to be understood. Often data-related activities are characterized by access to new data for which we have little or no prior knowledge of its inner structure and content. In these cases, we need to interactively *explore* the data first to gain insights and eventually be able to act upon the data contents. Interactive visual analysis is particularly useful in these cases where automatic analysis approaches fail and human capabilities need to be exploited and augmented.

Within this research scope Aviz focuses on five research themes:

- Methods to visualize and smoothly navigate through large datasets;
- Efficient analysis methods to reduce huge datasets to visualizable size;
- Visualization interaction using novel capabilities and modalities;
- Evaluation methods to assess the effectiveness of visualization and analysis methods and their usability;
- Engineering tools for building visual analytics systems that can access, search, visualize and analyze large datasets with smooth, interactive response.

2.2. Research Themes

Aviz's research on Visual Analytics is organized around five main Research Themes:

Methods to visualize and smoothly navigate through large data sets: Large data sets challenge current visualization and analysis methods. Understanding the structure of a graph with one million vertices is not just a matter of displaying the vertices on a screen and connecting them with lines. Current screens only have around two million pixels. Understanding a large graph requires both data reduction to visualize the whole and navigation techniques coupled with suitable representations to see the details. These representations, aggregation functions, navigation and interaction techniques must be chosen as a coordinated whole to be effective and fit the user's mental map.

Aviz designs new visualization representations and interactions to efficiently navigate and manipulate large data sets.

Efficient analysis methods to reduce huge data sets to visualizable size: Designing analysis components with interaction in mind has strong implications for both the algorithms and the processes they use. Some data reduction algorithms are suited to the principle of sampling, then extrapolating, assessing the quality and incrementally enhancing the computation: for example, all the linear reductions such as PCA, Factorial Analysis, and SVM, as well as general MDS and Self Organizing Maps. Aviz investigates the possible analysis processes according to the analyzed data types.

Visualization interaction using novel capabilities and modalities: The importance of interaction to Visualization and, in particular, to the interplay between interactivity and cognition is widely recognized. However, information visualization interactions have yet to take full advantage of these new possibilities in interaction technologies, as they largely still employ the traditional desktop, mouse, and keyboard setup of WIMP (Windows, Icons, Menus, and a Pointer) interfaces. At Aviz we investigate in particular interaction through tangible and touch-based interfaces to data.

Evaluation methods to assess their effectiveness and usability: For several reasons appropriate evaluation of visual analytics solutions is not trivial. First, visual analytics tools are often designed to be applicable to a variety of disciplines, for various different data sources, and data characteristics, and because of this variety it is hard to make general statements. Second, in visual analytics the specificity of humans, their work environment, and the data analysis tasks, form a multi-faceted evaluation context which is difficult to control and generalize. This means that recommendations for visual analytics solutions are never absolute, but depend on their context.

In our work we systematically connect evaluation approaches to visual analytics research—we strive to develop and use both novel as well as establish mixed-methods evaluation approaches to derive recommendations on the use of visual analytics tools and techniques. Aviz regularly published user studies of visual analytics and interaction techniques and takes part in dedicated workshops on evaluation.

Engineering tools: for building visual analytics systems that can access, search, visualize and analyze large data sets with smooth, interactive response.

Currently, databases, data analysis and visualization all use the concept of data tables made of tuples and linked by relations. However, databases are storage-oriented and do not describe the data types precisely. Analytical systems describe the data types precisely, but their data storage and computation model are not suited to interactive visualization. Visualization systems use in-memory data tables tailored for fast display and filtering, but their interactions with external analysis programs and databases are often slow.

Aviz seeks to merge three fields: databases, data analysis and visualization. Part of this merging involves using common abstractions and interoperable components. This is a long-term challenge, but it is a necessity because generic, loosely-coupled combinations will not achieve interactive performance.

Aviz's approach is holistic: these five themes are facets of building an analysis process optimized for discovery. All the systems and techniques Aviz designs support the process of understanding data and forming insights while minimizing disruptions during navigation and interaction.

3. Research Program

3.1. Scientific Foundations

The scientific foundations of Visual Analytics lie primarily in the domains of Visualization and Data Mining. Indirectly, it inherits from other established domains such as graphic design, Exploratory Data Analysis (EDA), statistics, Artificial Intelligence (AI), Human-Computer Interaction (HCI), and Psychology.

The use of graphic representation to understand abstract data is a goal Visual Analytics shares with Tukey's Exploratory Data Analysis (EDA) [47], graphic designers such as Bertin [36] and Tufte [46], and HCI researchers in the field of Information Visualization [35].

EDA is complementary to classical statistical analysis. Classical statistics starts from a *problem*, gathers *data*, designs a *model* and performs an *analysis* to reach a *conclusion* about whether the data follows the model. While EDA also starts with a problem and data, it is most useful *before* we have a model; rather, we perform visual analysis to discover what kind of model might apply to it. However, statistical validation is not always required with EDA; since often the results of visual analysis are sufficiently clear-cut that statistics are unnecessary.

Visual Analytics relies on a process similar to EDA, but expands its scope to include more sophisticated graphics and areas where considerable automated analysis is required before the visual analysis takes place. This richer data analysis has its roots in the domain of Data Mining, while the advanced graphics and interactive exploration techniques come from the scientific fields of Data Visualization and HCI, as well as the expertise of professions such as cartography and graphic designers who have long worked to create effective methods for graphically conveying information.

The books of the cartographer Bertin and the graphic designer Tufte are full of rules drawn from their experience about how the meaning of data can be best conveyed visually. Their purpose is to find effective visual representation that describe a data set but also (mainly for Bertin) to discover structure in the data by using the right mappings from abstract dimensions in the data to visual ones.

For the last 25 years, the field of Human-Computer Interaction (HCI) has also shown that interacting with visual representations of data in a tight perception-action loop improves the time and level of understanding of data sets. Information Visualization is the branch of HCI that has studied visual representations suitable to understanding and interaction methods suitable to navigating and drilling down on data. The scientific foundations of Information Visualization come from theories about perception, action and interaction.

Several theories of perception are related to information visualization such as the "Gestalt" principles, Gibson's theory of visual perception [40] and Triesman's "preattentive processing" theory [45]. We use them extensively but they only have a limited accuracy for predicting the effectiveness of novel visual representations in interactive settings.

Information Visualization emerged from HCI when researchers realized that interaction greatly enhanced the perception of visual representations.

To be effective, interaction should take place in an interactive loop faster than 100ms. For small data sets, it is not difficult to guarantee that analysis, visualization and interaction steps occur in this time, permitting smooth data analysis and navigation. For larger data sets, more computation should be performed to reduce the data size to a size that may be visualized effectively.

In 2002, we showed that the practical limit of InfoVis was on the order of 1 million items displayed on a screen [38]. Although screen technologies have improved rapidly since then, eventually we will be limited by the physiology of our vision system: about 20 millions receptor cells (rods and cones) on the retina. Another problem will be the limits of human visual attention, as suggested by our 2006 study on change blindness in large and multiple displays [37]. Therefore, visualization alone cannot let us understand very large data sets. Other techniques such as aggregation or sampling must be used to reduce the visual complexity of the data to the scale of human perception.

Abstracting data to reduce its size to what humans can understand is the goal of Data Mining research. It uses data analysis and machine learning techniques. The scientific foundations of these techniques revolve around the idea of finding a good model for the data. Unfortunately, the more sophisticated techniques for finding models are complex, and the algorithms can take a long time to run, making them unsuitable for an interactive environment. Furthermore, some models are too complex for humans to understand; so the results of data mining can be difficult or impossible to understand directly.

Unlike pure Data Mining systems, a Visual Analytics system provides analysis algorithms and processes compatible with human perception and understandable to human cognition. The analysis should provide understandable results quickly, even if they are not ideal. Instead of running to a predefined threshold, algorithms and programs should be designed to allow trading speed for quality and show the tradeoffs interactively. This is not a temporary requirement: it will be with us even when computers are much faster, because good quality algorithms are at least quadratic in time (e.g. hierarchical clustering methods). Visual Analytics systems need different algorithms for different phases of the work that can trade speed for quality in an understandable way.

Designing novel interaction and visualization techniques to explore huge data sets is an important goal and requires solving hard problems, but how can we assess whether or not our techniques and systems provide real improvements? Without this answer, we cannot know if we are heading in the right direction. This is why we have been actively involved in the design of evaluation methods for information visualization [44], [43], [41], [42], [39]. For more complex systems, other methods are required. For these we want to focus on longitudinal evaluation methods while still trying to improve controlled experiments.

3.2. Innovation

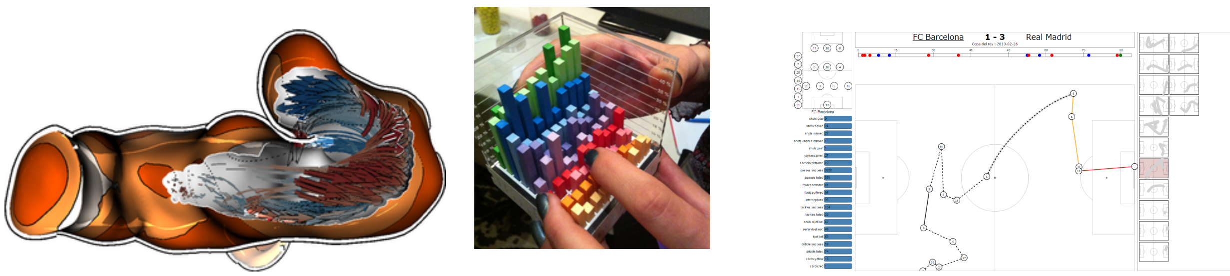


Figure 1. Example novel visualization techniques and tools developed by the team. Left: a non-photorealistic rendering technique that visualizes blood flow and vessel thickness. Middle:a physical visualization showing economic indicators for several countries, right: SoccerStories a tool for visualizing soccer games.

We design novel visualization and interaction techniques (see, for example, Figure 1). Many of these techniques are also evaluated throughout the course of their respective research projects. We cover application domains such as sports analysis, digital humanities, fluid simulations, and biology. A focus of Aviz' work is the improvement of graph visualization and interaction with graphs. We further develop individual techniques for the design of tabular visualizations and different types of data charts. Another focus is the use of animation as a transition aid between different views of the data. We are also interested in applying techniques from illustrative visualization to visual representations and applications in information visualization as well as scientific visualization.

3.3. Evaluation Methods

Evaluation methods are required to assess the effectiveness and usability of visualization and analysis methods. Aviz typically uses traditional HCI evaluation methods, either quantitative (measuring speed and errors) or qualitative (understanding users tasks and activities). Moreover, Aviz is also contributing to the improvement of evaluation methods by reporting on the best practices in the field, by co-organizing workshops (BELIV 2010, 2012, 2014, 2016) to exchange on novel evaluation methods, by improving our ways of reporting, interpreting and communicating statistical results, and by applying novel methodologies, for example to assess visualization literacy.

3.4. Software Infrastructures

We want to understand the requirements that software and hardware architectures should provide to support exploratory analysis of large amounts of data. So far, “big data” has been focusing on issues related to storage management and predictive analysis: applying a well-known set of operations on large amounts of data. Visual Analytics is about exploration of data, with sometimes little knowledge of its structure or properties. Therefore, interactive exploration and analysis is needed to build knowledge and apply appropriate analyses; this knowledge and appropriateness is supported by visualizations. However, applying analytical operations on large data implies long-lasting computations, incompatible with interactions, and generates large amounts of results, impossible to visualize directly without aggregation or sampling. Visual Analytics has started to tackle these problems for specific applications but not in a general manner, leading to fragmentation of results and difficulties to reuse techniques from one application to the other. We are interested in abstracting-out the issues and finding general architectural models, patterns, and frameworks to address the Visual Analytics challenge in more generic ways.

3.5. Emerging Technologies

Figure 2. Example emerging technology solutions developed by the team for multi-display environments, wall displays, and token-based visualization.

We want to empower humans to make use of data using different types of display media and to enhance how they can understand and visually and interactively explore information. This includes novel display equipment and accompanying input techniques. The Aviz team specifically focuses on the exploration of the use of large displays in visualization contexts as well as emerging physical and tangible visualizations. In terms of interaction modalities our work focuses on using touch and tangible interaction. Aviz participates to the Digiscope project that funds 11 wall-size displays at multiple places in the Paris area (see http://www.digiscope.fr), connected by telepresence equipment and a Fablab for creating devices. Aviz is in charge of creating and managing the Fablab, uses it to create physical visualizations, and is also using the local wall-size display (called WILD) to explore visualization on large screens. The team also investigates the perceptual, motor and cognitive implications of using such technologies for visualization.

3.6. Psychology

More cross-fertilization is needed between psychology and information visualization. The only key difference lies in their ultimate objective: understanding the human mind vs. helping to develop better tools. We focus on understanding and using findings from psychology to inform new tools for information visualization. In many cases, our work also extends previous work in psychology. Our approach to the psychology of information visualization is largely holistic and helps bridge gaps between perception, action and cognition in the context of information visualization. Our focus includes the perception of charts in general, perception in large display environments, collaboration, perception of animations, how action can support perception and cognition, and judgment under uncertainty.

4. Highlights of the Year

4.1. Highlights of the Year

4.1.1. Awards

BEST PAPERS AWARDS :

[] **IEEE Transactions on Visualization and Computer Graphics**. J. ZHAO, M. GLUECK, P. ISENBERG, F. CHEVALIER, A. KHAN.

[] **IEEE Transactions on Visualization and Computer Graphics**. E. DIMARA, A. BEZERIANOS, P. DRAGICEVIC.

5. New Software and Platforms

5.1. Cartolabe

KEYWORD: Information visualization

FUNCTIONAL DESCRIPTION: The goal of Cartolabe is to build a visual map representing the scientific activity of an institution/university/domain from published articles and reports. Using the HAL Database and building upon the AnHALytics processing chain, Cartolabe provides the user with a map of the thematics, authors and articles and their dynamics along time. ML techniques are used for dimensionality reduction, cluster and topics identification, visualisation techniques are used for a scalable 2D representation of the results.

NEWS OF THE YEAR: Improvement of the graphical interface

- Contact: Philippe Caillou
- URL: http://cartolabe.lri.fr/

5.2. BitConduite

BitConduite Bitcoin explorer

KEYWORDS: Data visualization - Clustering - Financial analysis - Cryptocurrency

FUNCTIONAL DESCRIPTION: BitConduite is a web-based visual tool that allows for a high level explorative analysis of the Bitcoin blockchain. It offers a data transformation back end that gives us an entity-based access to the blockchain data and a visualization front end that supports a novel high-level view on transactions over time. In particular, it facilitates the exploration of activity through filtering and clustering interactions. This gives analysts a new perspective on the data stored on the blockchain.

- Contact: Petra Isenberg

6. New Results

6.1. HCI Requirements for Progressive Data Analysis

Participants: Jean-Daniel Fekete [correspondant], Sriram Karthik Badam, Niklas Elmqvist.

Progressive visual analytics (PVA) has emerged in recent years to manage the latency of data analysis systems. When analysis is performed progressively, rough estimates of the results are generated quickly and are then improved over time. Analysts can therefore monitor the progression of the results, steer the analysis algorithms, and make early decisions if the estimates provide a convincing picture. In this article, we describe interface design guidelines for helping users understand progressively updating results and make early decisions based on progressive estimates. To illustrate our ideas, we present a prototype PVA tool called INSIGHTSFEED for exploring Twitter data at scale. As validation, we investigate the tradeoffs of our tool when exploring a Twitter dataset in a user study. We report the usage patterns in making early decisions using the user interface, guiding computational methods, and exploring different subsets of the dataset, compared to sequential analysis without progression.

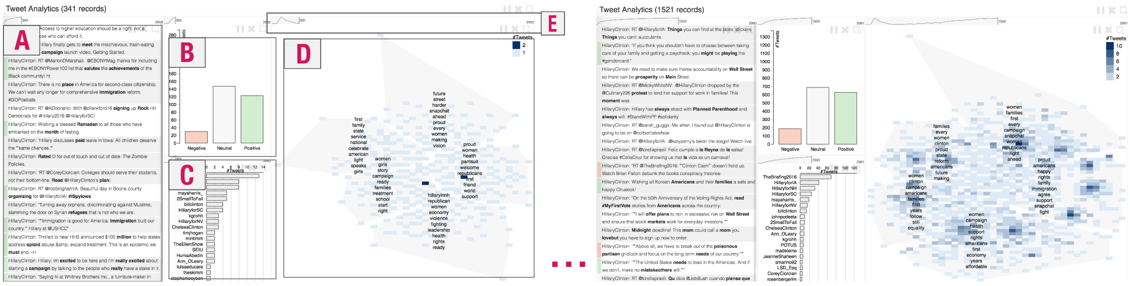


Figure 3. The InsightsFeed tool for progressive visual analytics of Twitter (left): (A) a list of tweets, (B) a sentiment chart, (C) user popularity chart, (D) a map from a 2D projection of tweets with important keywords highlighted in each region, and (E) feedback and controls over the progression and computations. (Right) The interface is progressively updated when more data is processed.

More on the project Web page: ProgressiveDataAnalysis.

6.2. Embedded Data Representations

Participants: Wesley Willett, Yvonne Jansen, Pierre Dragicevic [correspondant].

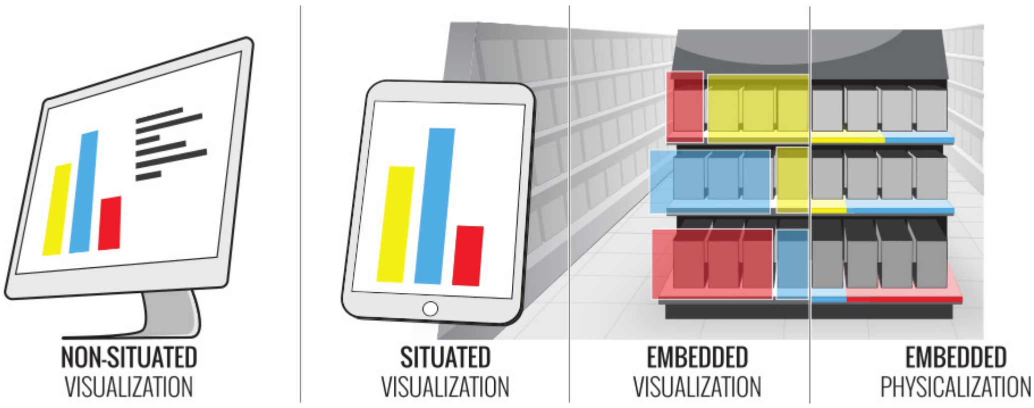


Figure 4. From left to right: A desktop setting with non-situated visualization. A situated visualization of the same data on a tablet in the store itself. An embedded visualization overlays the data on top of individual products as a heat map. An embedded physicalization displays data by changing properties of the shelves themselves.

We introduced *embedded data representations*, the use of visual and physical representations of data that are deeply integrated with the physical spaces, objects, and entities to which the data refers. Technologies like lightweight wireless displays, mixed reality hardware, and autonomous vehicles are making it increasingly easier to display data in-context. While researchers and artists have already begun to create embedded data representations, the benefits, trade-offs, and even the language necessary to describe and compare these approaches remain unexplored. In our paper [17], we formalized the notion of physical data referents – the real-world entities and spaces to which data corresponds – and examined the relationship between referents and the visual and physical representations of their data. We differentiated situated representations, which display data in proximity to data referents, and embedded representations, which display data so that it spatially coincides with data referents. Drawing on examples from visualization, ubiquitous computing, and art, we explored the role of spatial indirection, scale, and interaction for embedded representations. We also examined the tradeoffs between non-situated, situated, and embedded data displays, including both visualizations and physicalizations. Based on our observations, we identified a variety of design challenges for embedded data representation, and suggested opportunities for future research and applications.

More on the project Web page: yvonnejansen.me/embedded.

6.3. Blinded with Science or Informed by Charts? A Replication Study

Participants: Pierre Dragicevic [correspondant], Yvonne Jansen.

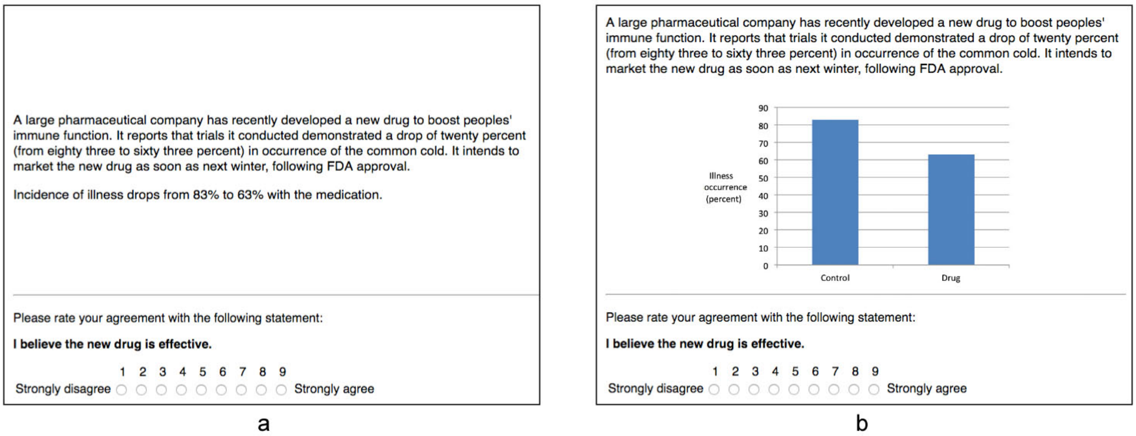


Figure 5. First page of our second experiment, replicating experiment 2 from Tal and Wansink. (a) no-chart condition, with an extra sentence repeating the two quantities with numerals; (b) chart condition: the extra sentence is replaced with a bar chart.

We provided a reappraisal of Tal and Wansink's study "Blinded with Science", where seemingly trivial charts were shown to increase belief in drug efficacy, presumably because charts are associated with science. Through a series of four replications conducted on two crowdsourcing platforms, we investigated an alternative explanation, namely, that the charts allowed participants to better assess the drug's efficacy. Considered together, our experiments suggested that the chart seems to have indeed promoted understanding, although the effect is likely very small. Meanwhile, we were unable to replicate the original study's findings, as text with chart appeared to be no more persuasive – and sometimes less persuasive – than text alone. This suggested that the effect may not be as robust as claimed and may need specific conditions to be reproduced. Regardless, within our experimental settings and considering our study as a whole (N = 623), the chart's contribution to understanding was clearly larger than its contribution to persuasion.

The main lesson from our study is that with charts, the peripheral route of persuasion cannot be studied independently from the central route: in order to establish that a chart biases judgment, it is necessary to also rigorously establish that it does not aid comprehension. Our replication also opens many relevant questions for infovis. Are charts really associated with science? More generally, what associations do charts or visualizations trigger depending on their visual design? When exactly is a chart trivial? Two arguments against minimalistic charts is that they take up space and they break the flow of the text. How do word-scale visualizations change these trade-offs?

Experimental material can be downloaded here: www.aviz.fr/blinded.

6.4. Vispubdata

Participants: Petra Isenberg [correspondant], Florian Heimerl, Steffen Koch, Tobias Isenberg, Panpan Xu, Charles Stolper, Michael Sedlmair, Torsten Möller, John Stasko.

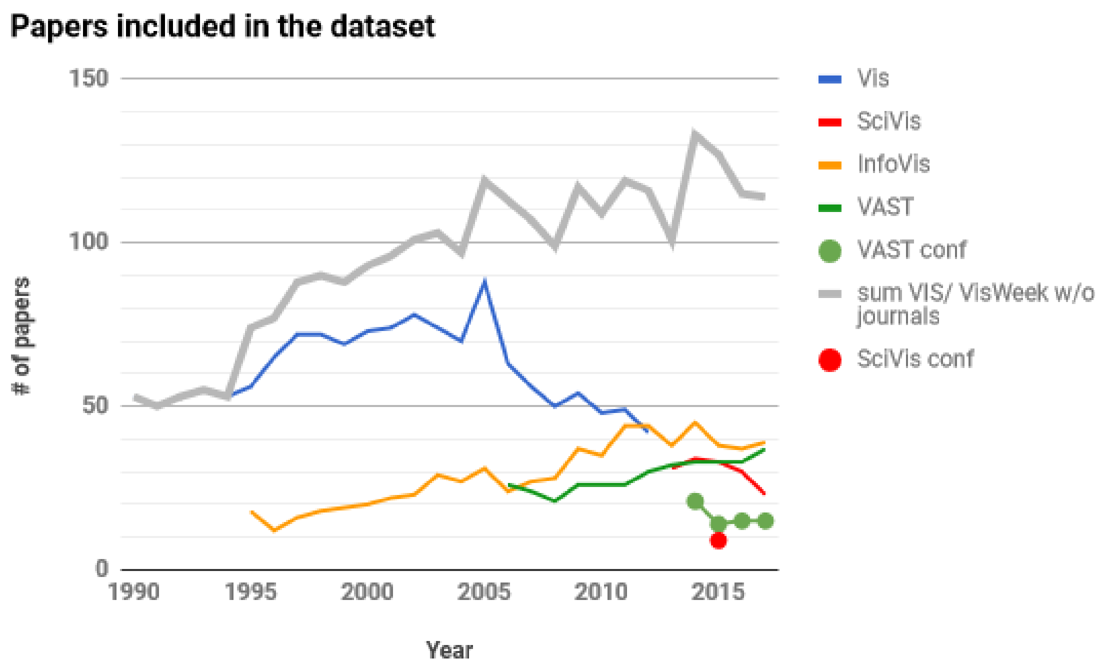


Figure 6. Overview of the files included in the dataset.

We have created and keep maintaining a dataset with information about every paper that has appeared at the IEEE Visualization (VIS) set of conferences: InfoVis, SciVis, VAST, and Vis. The information about each paper includes its title, abstract, authors, and citations to other papers in the conference series, among many other attributes. This data is meant to be useful to the broad data visualization community to help understand the evolution of the field and as an example document collection for text data visualization research.

6.5. An Exploratory Study of Word-Scale Graphics in Data-Rich Text Documents

Participants: Pascal Goffin, Jeremy Boy, Wesley Willett, Petra Isenberg [correspondant].

We contribute an investigation of the design and function of word-scale graphics and visualizations embedded in text documents. Word-scale graphics include both data-driven representations such as word-scale visualizations and sparklines, and non-data-driven visual marks. Their design, function, and use has so far received little research attention. We conducted an open ended exploratory study with 9 graphic designers. The study resulted in a rich collection of different types of graphics, data provenance, and relationships between text, graphics, and data. Based on this corpus, we derived a systematic overview of word-scale graphic designs, and examine how designers used them. We also discussed the designers' goals in creating their graphics, and

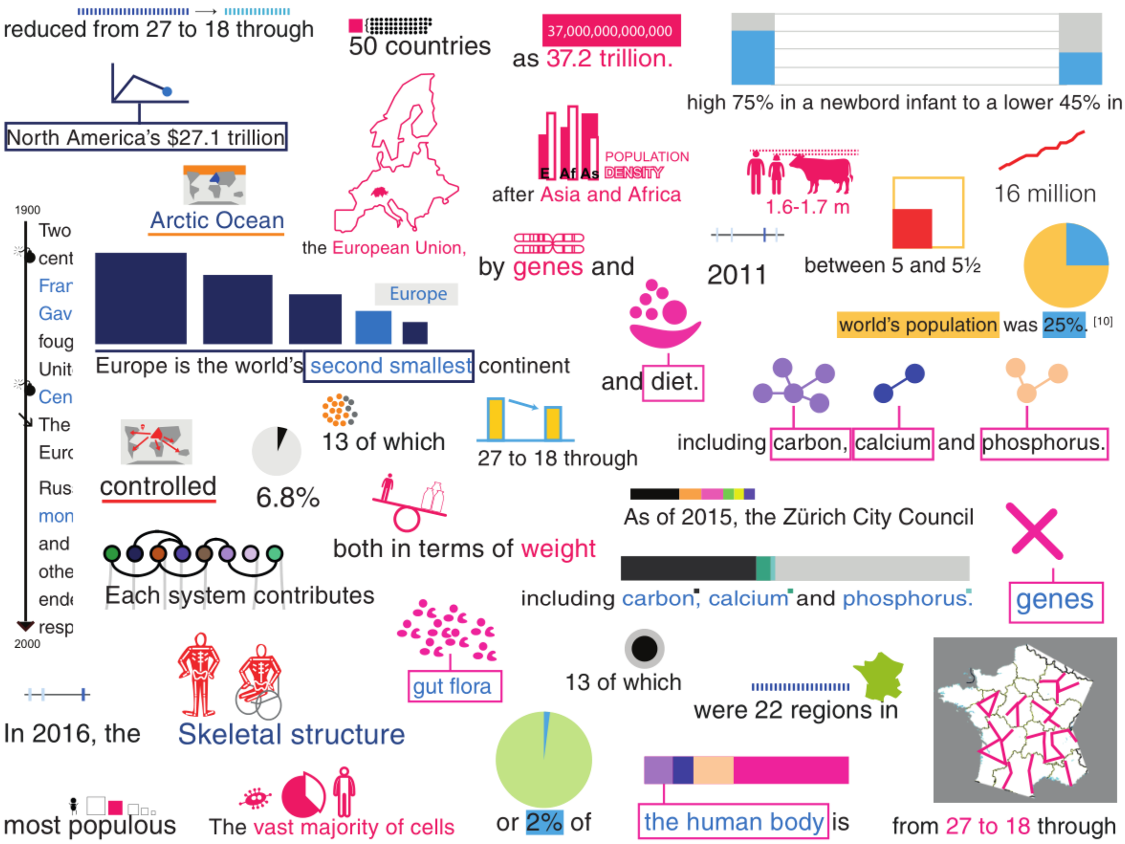


Figure 7. Overview of the word-scale visualizations created in our study.

characterized how they used word-scale graphics to visualize data, add emphasis, and create alternative narratives. Building on these examples, we discuss implications for the design of authoring tools for word-scale graphics and visualizations, and explore how new authoring environments could make it easier for designers to integrate them into documents.

6.6. Hybrid Tactile/Tangible Interaction for 3D Data Exploration

Participants: Lonni Besançon [correspondant], Paul Issartel, Mehdi Ammi, Tobias Isenberg.

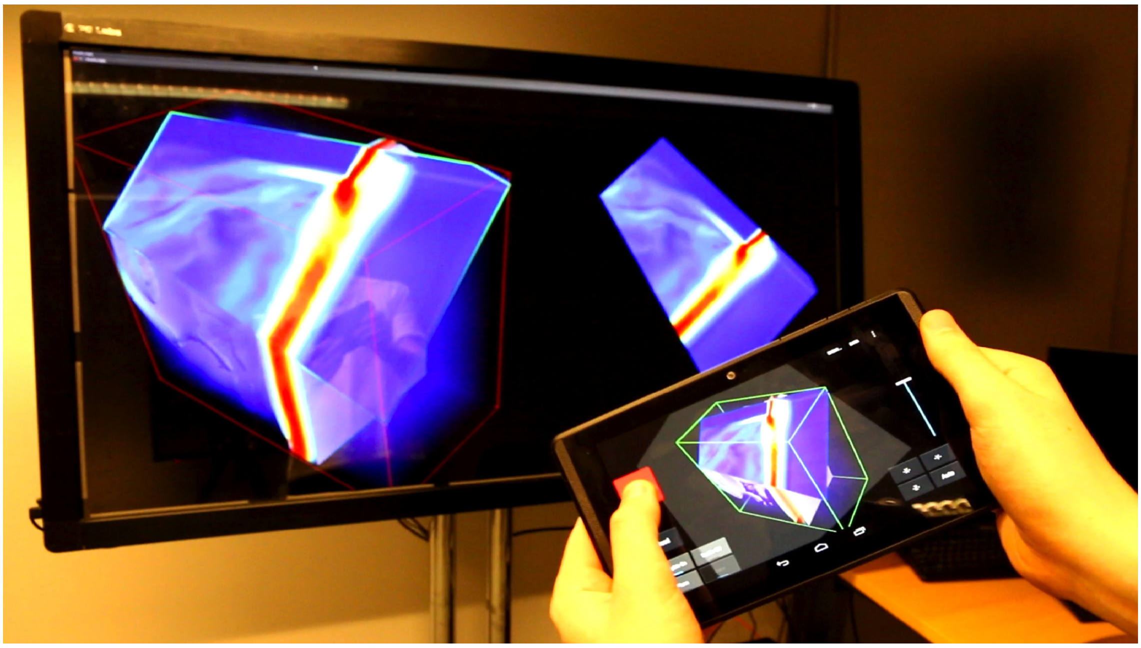


Figure 8. Picture of the hybrid interaction system.

We present the design and evaluation of an interface that combines tactile and tangible paradigms for 3D visualization. While studies have demonstrated that both tactile and tangible input can be efficient for a subset of 3D manipulation tasks, we reflect here on the possibility to combine the two complementary input types. Based on a field study and follow-up interviews, we present a conceptual framework of the use of these different interaction modalities for visualization both separately and combined—focusing on free exploration as well as precise control. We present a prototypical application of a subset of these combined mappings for fluid dynamics data visualization using a portable, position-aware device which offers both tactile input and tangible sensing. We evaluate our approach with domain experts and report on their qualitative feedback.

More on the project Web page: http://lonni.besancon.pagesperso-orange.fr/Projects/HybridInteraction/HybridInteraction.html.

6.7. Pressure-Based Gain Factor Control for Mobile 3D Interaction using Locally-Coupled Devices

Participants: Lonni Besançon [correspondant], Mehdi Ammi, Tobias Isenberg.

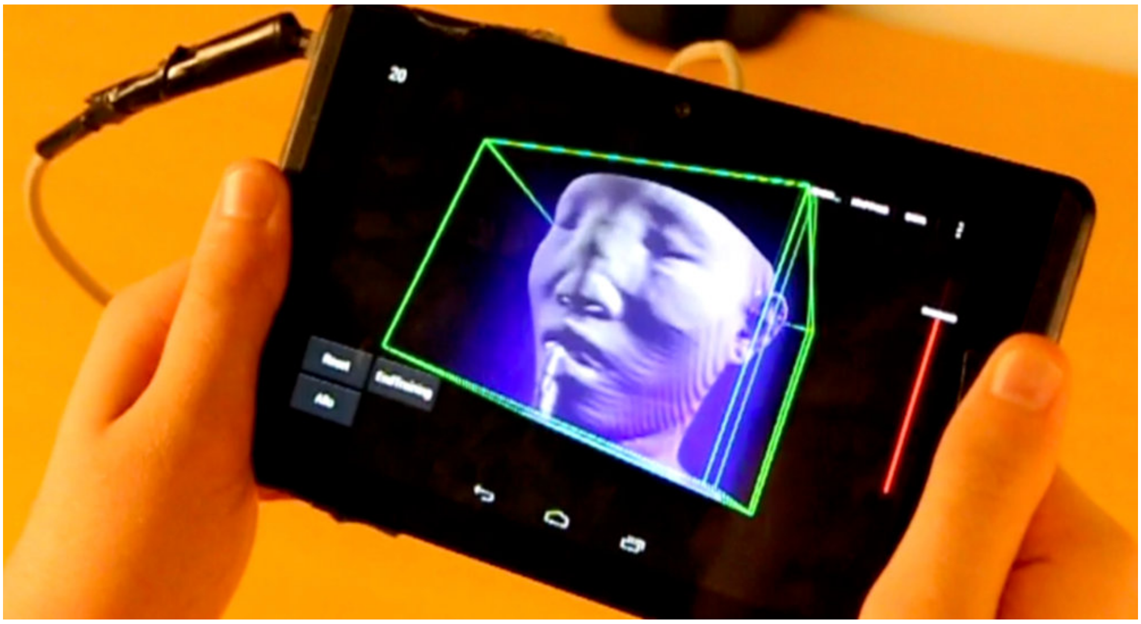


Figure 9. Picture of the pressure-based gain factor control prototype.

We present the design and evaluation of pressure-based interactive control of 3D navigation precision. Specifically, we examine the control of gain factors in tangible 3D interactions using locally-coupled mobile devices. By focusing on pressure as a separate input channel we can adjust gain factors independently from other input modalities used in 3D navigation, in particular for the exploration of 3D visualizations. We present two experiments. First, we determined that people strongly preferred higher pressures to be mapped to higher gain factors. Using this mapping, we compared pressure with rate control, velocity control, and slider-based control in a second study. Our results show that pressure-based gain control allows people to be more precise in the same amount of time compared to established input modalities. Pressure-based control was also clearly preferred by our participants. In summary, we demonstrate that pressure facilitates effective and efficient precision control for mobile 3D navigation.

More on the project Web page: http://lonni.besancon.pagesperso-orange.fr/Projects/Pressure/Pressure.html.

6.8. Pressure-Based Gain Factor Control for Mobile 3D Interaction using Locally-Coupled Devices

Participants: Lonni Besançon [correspondant], Mehdi Ammi, Tobias Isenberg.

We present the design and evaluation of pressure-based interactive control of 3D navigation precision. Specifically, we examine the control of gain factors in tangible 3D interactions using locally-coupled mobile devices. By focusing on pressure as a separate input channel we can adjust gain factors independently from other input modalities used in 3D navigation, in particular for the exploration of 3D visualizations. We present two experiments. First, we determined that people strongly preferred higher pressures to be mapped to higher gain factors. Using this mapping, we compared pressure with rate control, velocity control, and slider-based control in a second study. Our results show that pressure-based gain control allows people to be more precise in the same amount of time compared to established input modalities. Pressure-based control was also clearly preferred by our participants. In summary, we demonstrate that pressure facilitates effective and efficient precision control for mobile 3D navigation.

6.9. The Attraction Effect in Information Visualization

Participants: Evanthia Dimara [correspondant], Pierre Dragicevic, Anastasia Bezerianos.

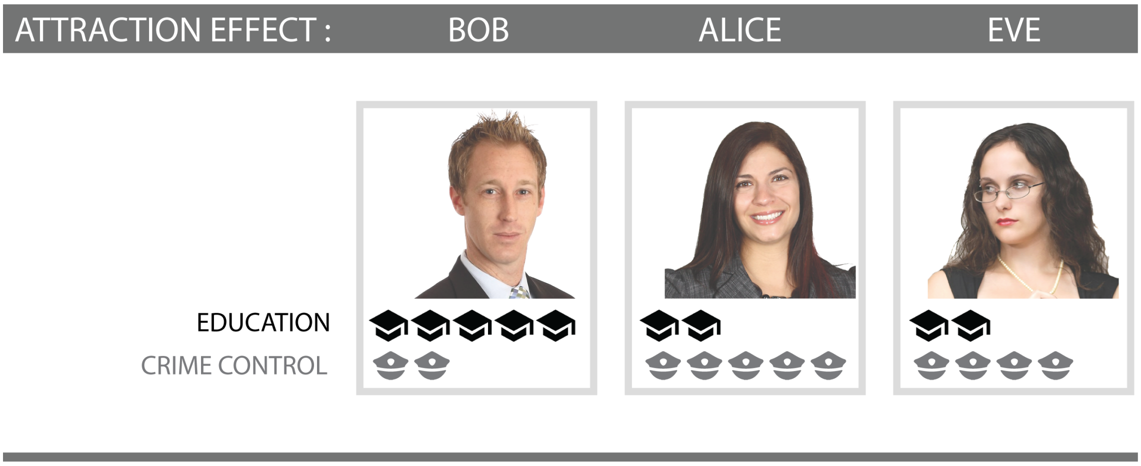


Figure 10. Example of an attraction effect in elections: Bob has an excellent education plan, while Alice is very strong in crime control. The addition of Eve, a candidate similar but slightly inferior to Alice, raises Alice's attractiveness as a candidate. This irrelevant option is called a decoy.

The attraction effect is a well-studied cognitive bias in decision making research, where one's choice between two alternatives is influenced by the presence of an irrelevant (dominated) third alternative. We examine whether this cognitive bias, so far only tested with three alternatives and simple presentation formats such as numerical tables, text and pictures, also appears in visualiza-tions. Since visualizations can be used to support decision making — e.g., when choosing a house to buy or an employee to hire — a systematic bias could have important implications. In a first crowdsource experiment, we indeed partially replicated the attraction effect with three alternatives presented as a numerical table, and observed similar effects when they were presented as a scatterplot. In a second experiment, we investigated if the effect extends to larger sets of alternatives, where the number of alternatives is too large for numerical tables to be practical. Our findings indicate that the bias persists for larger sets of alternatives presented as scatterplots. We discuss implications for future research on how to further study and possibly alleviate the attraction effect.

More on the project Web page: http://www.aviz.fr/decoy.

6.10. Narratives in Crowdsourced Evaluation of Visualizations: A Double-Edged Sword?

Participants: Evanthia Dimara [correspondant], Pierre Dragicevic, Anastasia Bezerianos.

We explore the effects of providing task context when evaluating visualization tools using crowdsourcing. We gave crowdworkers i) abstract information visualization tasks without any context, ii) tasks where we added semantics to the dataset, and iii) tasks with two types of backstory narratives: an analytic narrative and a decision-making narrative. Contrary to our expectations, we did not find evidence that adding data semantics increases accuracy, and further found that our backstory narratives can even decrease accuracy. Adding dataset semantics can however increase attention and provide subjective benefits in terms of confidence, perceived easiness, task enjoyability and perceived usefulness of the visualization. Nevertheless, our backstory narratives

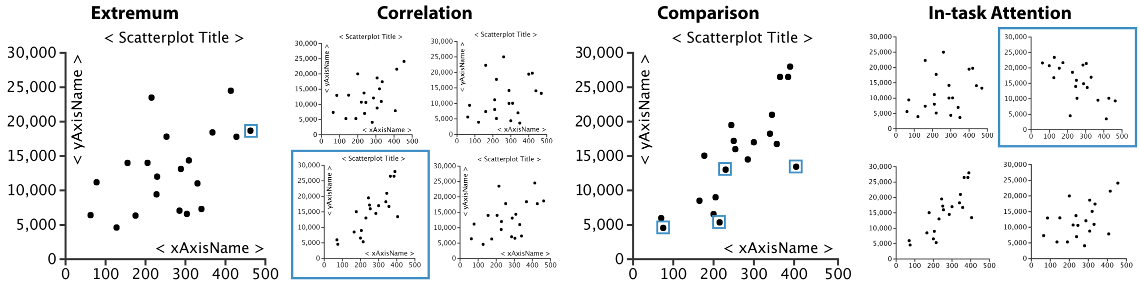


Figure 11. Stimuli used in each task (Ext, Cor and Com), and in the in-task attention test. Correct answers are annotated in blue. Axes were labeled (X,Y) for ABS, and (size m2, price ($)) in all other context conditions. The title was Diagram Z : Datapoints in ABS, and was Diagram Z : Houses in SEM (all tasks) and DM-NAR (Ext, Cor tasks). In all other conditions the title was Agency Z : Houses. Z was an integer (1, 2, 3, or 4) identifying the scatterplot.

did not appear to provide additional subjective benefits. These preliminary findings suggest that narratives may have complex and unanticipated effects, calling for more studies in this area.

More on the project Web page: http://www.aviz.fr/narratives.

6.11. Conceptual and Methodological Issues in Evaluating Multidimensional Visualizations for Decision Support.

Participants: Evanthia Dimara [correspondant], Pierre Dragicevic, Anastasia Bezerianos.

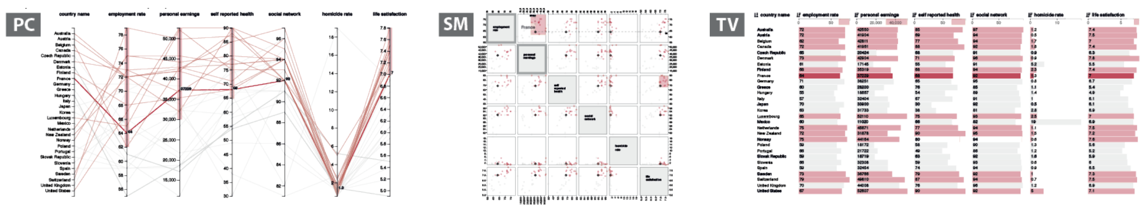


Figure 12. The visualizations we evaluated: Parallel Coordinates (PC), Scatterplot Matrix (SM) and Tabular Visualization (TV).

We explore how to rigorously evaluate multidimensional visualizations for their ability to support decision making. We first define multi-attribute choice tasks, a type of decision task commonly performed with such visualizations. We then identify which of the existing multidimensional visualizations are compatible with such tasks, and set out to evaluate three elementary visualizations: parallel coordinates, scatterplot matrices and tabular visualizations. Our method consists in first giving participants low-level analytic tasks, in order to ensure that they properly understood the visualizations and their interactions. Participants are then given multi-attribute choice tasks consisting of choosing holiday packages. We assess decision support through multiple objective and subjective metrics, including a decision accuracy metric based on the consistency between the choice made and self-reported preferences for attributes. We found the three visualizations to be comparable on most metrics, with a slight advantage for tabular visualizations. In particular, tabular visualizations allow participants to reach decisions faster. Thus, although decision time is typically not central in assessing decision support, it can be used as a tie-breaker when visualizations achieve similar decision accuracy. Our results also suggest that indirect methods for assessing choice confidence may allow to better distinguish between visualizations than direct ones. We finally discuss the limitations of our methods and directions for future work, such as the need for more sensitive metrics of decision support.

More on the project Web page: http://www.aviz.fr/dm.

7. Partnerships and Cooperations

7.1. European Initiatives

7.1.1. FP7 & H2020 Projects

7.1.1.1. CENDARI

Title: Collaborative EuropeaN Digital/Archival Infrastructure

Programm: FP7

Duration: February 2012 - January 2016

Coordinator: Trinity College - Dublin

Partners:

Consortium of European Research Libraries (United Kingdom)

Koninklijke Bibliotheek (Netherlands)

Fondazione Ezio Franceschini Onlus (Italy)

Freie Universitaet Berlin (Germany)

King's College London (United Kingdom)

"matematicki Institutnu, Beograd" (Serbia)

Narodni Knihovna Ceske Republiky (Czech Republic)

Societa Internazionale Per Lo Studio Del Medioevo Latino-S.I.S.M.E.L.Associazione (Italy)

The Provost, Fellows, Foundation Scholars & The Other Members of Board of The College of The Holy & Undivided Trinity of Queen Elizabeth Near Dublin (Ireland)

Georg-August-Universitaet Goettingen Stiftung Oeffentlichen Rechts (Germany)

The University of Birmingham (United Kingdom)

Universitaet Stuttgart (Germany)

Universita Degli Studi di Cassino E Del Lazio Meridionale (Italy)

Inria contact: L. Romary

'The Collaborative EuropeaN Digital Archive Infrastructure (CENDARI) will provide and facilitate access to existing archives and resources in Europe for the study of medieval and modern European history through the development of an 'enquiry environment'. This environment will increase access to records of historic importance across the European Research Area, creating a powerful new platform for accessing and investigating historical data in a transnational fashion overcoming the national and institutional data silos that now exist. It will leverage the power of the European infrastructure for Digital Humanities (DARIAH) bringing these technical experts together with leading historians and existing research infrastructures (archives, libraries and individual digital projects) within a programme of technical research informed by cutting edge reflection on the impact of the digital age on scholarly practice. The enquiry environment that is at the heart of this proposal will create new ways to discover meaning, a methodology not just of scale but of kind. It will create tools and workspaces that allow researchers to engage with large data sets via federated multilingual searches across heterogeneous resources while defining workflows enabling the creation of personalized research environments, shared research and teaching spaces, and annotation trails, amongst other features. This will be facilitated by multilingual authority lists of named entities (people, places, events) that will harness user involvement to add intelligence to the system. Moreover, it will develop new visual paradigms for the exploration of patterns generated by the system, from knowledge transfer and dissemination, to language usage and shifts, to the advancement and diffusion of ideas.'

7.2. International Initiatives

7.2.1. Informal International Partners

- University of Calgary. Pierre Dragicevic and Petra Isenberg collaborate with Wesley Willett on situated data visualization.
- University of Washington, Chicago University and University of Zurich. Pierre Dragicevic collaborates with Matthew Kay, Steve Haroz and Chat Wacharamanotham on transparent statistical reporting and efficient statistical communication
- Stanford University. Pierre Dragicevic and Jean-Daniel Fekete collaborate with Sean Follmer on swarm user interfaces.
- Chicago University and University of Maryland, Evanthia Dimara and Pierre Dragicevic collaborate with Steven Franconeri and Catherine Plaisant on a taxonomy of cognitive biases.

7.3. International Research Visitors

7.3.1. Visits of International Scientists

- Catherine Plaisant (June–July): Invited professor from University of Maryland, USA. Invited through a DigiCosme grant, Catherine Plaisant has spent two months with Aviz. We have launched two research projects, one on hypergraph visualization and one on tracing users to understand their use of visualization. Catherine Plaisant has interacted with all of the Aviz students and post-doctoral fellows, as well as with the permanent researchers.
- Paolo Buono, from University of Bari, Italy. Paolo Buono has spent two months with Aviz working on the visualization of dynamic networks. He has collaborated with Paoa Valdivia, Catherine Plaisant, and Jean-Daniel Fekete for that project. He has also interacted with all the members of Aviz.

7.3.1.1. Internships

- Jaemin Jo (March–April): intern from Seoul National University, Korea. Worked on converting a KNN algorithm into a progressive form.
- Nicola Pezzotti (April–May): intern from University of Delft, The Netherlands. Worked on data structures and algorithms for managing very large (out of core) datasets in the context of progressive algorithms.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events Organisation

8.1.1.1. Member of the Organizing Committees

- Jean-Daniel Fekete organized the Doctoral Colloquium (Doctoriales) for the Journées Visu 2017
- Tobias Isenberg was Posters co-chair for EuroVis 2017
- Pierre Dragicevic co-organized the CHI 2017 workshop on Moving Transparent Statistics Forward @CHI 2017.
- Petra Isenberg was Posters co-chair for VIS 2017
- Evanthia Dimara organized the workshop "DECISIVe 2017: Dealing with Cognitive Biases in Visualisations" at IEEE VIS'2017 (http://decisive-workshop.dbvis.de/).

8.1.2. Scientific Events Selection

8.1.2.1. Member of the Conference Program Committees

- Jean-Daniel Fekete was a member of the program committee for IEEE VIS 2017.
- Jean-Daniel Fekete was a member of the test of time award committee for IEEE InfoVis 2017.
- Jean-Daniel Fekete was a member of the program committee for EuroVis 2017.
- Tobias Isenberg was a member of the program committee for ACM/Eurographics Expressive 2017.
- Tobias Isenberg was a member of the program committee for ACM SUI 2017.
- Tobias Isenberg was a member of the program committee for IEEE VISAP 2017.
- Tobias Isenberg was a member of the program committee for IEEE InfoVis 2017.
- Tobias Isenberg was a member of the program committee for EuroVis 2017, full & short papers.
- Tobias Isenberg was a member of the program committee for Eurographics 2017.
- Tobias Isenberg was a member of the program committee for SCCG 2017.
- Tobias Isenberg was a member of the program committee for IEEE 3DUI 2017.
- Pierre Dragicevic was a member of the program committee for IEEE VIS 2017.
- Pierre Dragicevic was a member of the program committee for IHM 2017.
- Pierre Dragicevic was a member of the program committee for the DECISIVE 2017 workshop.
- Petra Isenberg was a member of the program committee for IEEE InfoVis 2017.
- Petra Isenberg was a member of the program committee for EuroVA 2017.
- Petra Isenberg was a member of the short paper committee for EuroVis 2017.
- Petra Isenberg was a member of the program committee for IEEE PacificVis 2017.
- Tanja Blascheck was a member of the Program Committee – NIER and Tool Tracks for VISSOFT 2017.

8.1.2.2. Reviewer

- Jean-Daniel Fekete reviewed for CHI, EuroVis, InfoVis.
- Tobias Isenberg reviewed for InfoVis, SciVis, EuroVis, SIGGRAPH, Eurographics, CHI, VR, SCCG, 3DUI, ISS, Graphics Interface, PacificVis, Expressive, SUI, VISAP.
- Pierre Dragicevic reviewed for CHI, VIS, UIST, Interact, IHM.
- Petra Isenberg reviewed for EuroVA, EuroVis, Interact, PacificVis, InfoVis.
- Christoph Kinkeldey reviewed for ACM INWUT.
- Tanja Blascheck reviewed for EuroVis, PacificVis, VISSOFT (IEEE Working Conference on Software Visualization), VRST (ACM Symposium on Virtual Reality Software and Technology).
- Lonni Besançon reviewed for 3DUI, CHI, ICMI, IHM, ISS, OzCHI, VR, VRST.

8.1.3. Journal

8.1.3.1. Member of the Editorial Boards

- Tobias Isenberg is member of the editorial board of Elsevier's Computers & Graphics journal.

- Pierre Dragicevic is member of the editorial board of the Human-centric Computing and Information Sciences (HCIS) journal.

8.1.3.2. Reviewer - Reviewing Activities

- Jean-Daniel Fekete reviewed for TVCG.
- Tobias Isenberg reviewed for TVCG.
- Pierre Dragicevic reviewed for TVCG.
- Petra Isenberg reviewed for TVCG, InfoVis Journal.
- Christoph Kinkeldey reviewed for TVCG, IEEE CG&A, IJGIS, ISPRS IJGI.
- Tanja Blascheck reviewed for JEMR (Journal of Eye Movement Research), Pattern Recognition Journal, TVCG.
- Lonni Besançon reviewed for IEEE Consumer Electronic.

8.1.4. Invited Talks

- Jean-Daniel Fekete: Seminar Dataviz : idée, méthode et perception Visualisation d'information : des principes au passage à l'échelle, INED, Paris, December 4, 2017
- Jean-Daniel Fekete: Keynote EA Artificial Evolution 2017, Progressive Data Analysis: a new computation paradigm for scalability in exploratory data analysis. October 25, 2017
- Jean-Daniel Fekete: Keynote ChinaVis 2017, Qingdao, China, Progressive Data Analysis: a new computation paradigm for scalability in exploratory data analysis , July 17, 2017
- Jean-Daniel Fekete: Invited Talk China-Germany Workshop, Qingdao, China, Visualization for the People, July 16, 2017
- Jean-Daniel Fekete: Keynote WSOM+ 2017, LORIA, Nancy, France, Visualization of Complex Networks, June 28th, 2017
- Jean-Daniel Fekete: Keynote VISIGRAPP 2017, Porto, Portugal, Understanding Complex Networks, Feb. 27, 2017
- Jean-Daniel Fekete: Invited Talk, Univ. of Konstanz and Univ. of Stuttgart, ProgressiVis: A New Language Paradigm for Scalability in Exploratory Analytics, Feb. 9, 2017
- Jean-Daniel Fekete: Keynote Speech, Univ. of Grenoble, Visualizing [Dense, Dynamic, Complex] Networks, Feb. 2, 2017
- Jean-Daniel Fekete: Open Data Conference, La Rochelle, Panel on Open Data and Visualization, Jan. 17, 2017
- Tobias Isenberg: Invited talk at TU Wien, Austria, December 22: "Interactive Navigation and Selection using Tactile and Tangible Inputs for 3D Data Exploration"
- Tobias Isenberg: Invited talk at Visualization research group, University of Bergen, Norway, April 7: "Abstraction in Non-Photorealistic Rendering and Illustrative Visualizations."
- Tobias Isenberg: Invited talk at Dept. of Computer Science, University of Bergen, Norway, April 6: "Tactile Navigation and Selection for 3D Data Exploration."
- Tobias Isenberg: Invited talk at Connecting The Dots/Intelligent Trackers (CTD/WIT 2017), Orsay, France, March 7: "Abstraction in Scientific Data Visualization: Application to Brain Connectivity and Structural Biology."
- Pierre Dragicevic: "Statistical Dances: Why No Statistical Analysis is Reliable and What To Do About It". Séminaires Recherche Reproductible, LIG, Grenoble. 22 June 2017.
- Petra Isenberg: "Evaluation in Visualization: A closer look at current practices, issues, and perspectives". University of Stuttgart, Visualization Seminar Series
- Petra Isenberg: "Evaluation in Visualization: A closer look at current practices, issues, and perspectives". Technical University of Vienna, Visualization Seminar Series

- Tanja Blascheck: “Evaluating Interactive Visualizations using Eye Tracking, Interaction Logs, and Think-Aloud Protocols”. Inria, Aviz, Saclay, France, 3 March 2017.
- Kuno Kurzhals and Tanja Blascheck: “Visual Analysis of Eye Tracking Data”. Popakademie, Mannheim, Germany, 3 July 2017.
- Tanja Blascheck: “ Understanding Interactive Visualizations: Leveraging Eye Movements and Visual Analytics”. Hochschule Karlsruhe, Karlsruhe, Germany, 31 August 2017.
- Evanthia Dimara: “The Attraction Effect in Information Visualization”. Invited talk at University of Athens, 28 September 2017.
- Evanthia Dimara: “Biais cognitifs dans la visualisation d’information: implications pour l’évaluation”. Talk at Visu 2017.

8.1.5. Leadership within the Scientific Community

- Jean-Daniel Fekete is a member of the Steering Committee of EuroVis (Eurographics WG on Data Visualization).
- Jean-Daniel Fekete is a member of the Steering Committee of the IEEE Information Visualization Conference.
- Jean-Daniel Fekete is a member of the Eurographics Publication Board.
- Tobias Isenberg is a member of the Executive Committee of the Visualization and Computer Graphics Technical Committee of the IEEE Computer Society and serves as Publications Chair
- Tobias Isenberg is a member of the Steering Committee of Expressive (Joined Symposium on Computational Aesthetics, Sketch-Based Interfaces & Modeling, and Non-Photorealistic Animation & Rendering)

8.1.6. Scientific Expertise

- Tobias Isenberg reviewed for NSERC.
- Pierre Dragicevic reviewed for NSERC.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

- “Information Visualization” taught by Jean-Daniel Fekete at Polytechnique, Data Science, Saclay, France, Dec. 2017
- “Big Data Visual Analytics” seminar taught by Jean-Daniel Fekete at Centrale Supelec, Decision Support and Business of Intelligence, Nov. 28th, Saclay, France, 2017
- “Information Visualization and Visual Analytics” taught by Jean-Daniel Fekete at ISCD Summer School, Roscoff, France, August 2017
- “Photorealistic Rendering” taught by Tobias Isenberg at Polytech Paris-Sud and Université Paris-Saclay, France
- “Introduction to Computer Graphics” taught by Tobias Isenberg at Polytech Paris-Sud, France
- “Non-Photorealistic Rendering” taught by Tobias Isenberg at the University of Granada, Spain
- “Introduction to Statistics” guest lecture by Pierre Dragicevic for the Visual Analytics Master Course at École Centrale, Paris.
- “Interactive Information Visualization” taught by Petra Isenberg at Université Paris Sud
- “Visual Analytics” taught by Petra Isenberg at CentraleSupelec

8.2.2. Supervision

- PhD: Evanthia Dimara, Information Visualization for Decision Making, Université Paris-Sud, 2014, Pierre Dragicevic and Anastasia Bezerianos. Defended on Nov 30.

- PhD: Lonni Besançon, An Interaction Continuum for Scientific Visualization, Université Paris-Sud, 2014, Tobias Isenberg. Defended on December 14.
- PhD: Paola Llerena Valdivia, Wavelet-based analysis of time-varying data on graphs, University of São Paulo - USP, 2017, Jean-Daniel Fekete, Luis Gustavo Nonato.
- PhD: Marc Barnabé, Multiscale reconstruction of microbial ecosystems using semi-supervised machine learning, Université Paris-Sud, 2017, Jean-Daniel Fekete, Evelyne Lutton, INRA.
- PhD in progress: Xiyao Wang, Augmented Reality Environments for the Interactive Exploration of 3D Data, Univ. Paris-Sud; 2017, Tobias Isenberg
- PhD in progress: Sarkis Halladjian, Spatially Integrated Abstraction of Genetic Molecules, Univ. Paris-Sud; 2017, Tobias Isenberg
- PhD in progress: Haichao Miao, Visual Abstraction and Modeling for DNA Nanotechnology, TU Wien, Austria, 2016, Tobias Isenberg

8.2.3. Juries

- Jean-Daniel Fekete: Member of the PhD committee of Dr. Arnaud Prouzeau.
- Jean-Daniel Fekete: Member of the PhD committee of Dr. Antoine Lhuillier.
- Jean-Daniel Fekete: Member of the PhD committee of Dr. Alexandre Perrot.
- Jean-Daniel Fekete: Member of the PhD committee of Dr. Nicolas Médoc.
- Jean-Daniel Fekete: Member of the HdR committee of Dr. Jean-Philippe Cointet.
- Jean-Daniel Fekete: Member of the HdR committee of Dr. Gilles Bailly.
- Tobias Isenberg: Member of the PhD committee of Dr. Paul Issartel.
- Pierre Dragicevic: Hiring committee for tenure-track assistant Professor of Information Visualization, University of Copenhagen.
- Pierre Dragicevic: Member of the Commission Consultative de Spécialistes de l'Université Paris-Sud (CCSU).
- Pierre Dragicevic: Mid-term PhD evaluation committee of Michael Wessely.
- Pierre Dragicevic: Mid-term PhD evaluation committee of Abby Liu.
- Pierre Dragicevic: Reviewer for Ruoqi He's M2 internship.
- Petra Isenberg: Jury Inria Starting Research Positions

8.3. Popularization

- Jean-Daniel Fekete contributed to the DataVis section of the Terra Data exhibition at La Cité des Sciences, April 4th 2017 to January 7th 2018
- Jean-Daniel Fekete contributed to the exhibition on Jacques Bertin at EHESS Paris, November 14th to December 15th 2017
- Jean-Daniel Fekete and Charles Perin organized a workshop on data visualization for the "retrospective Jacques Bertin" at EHESS Paris on November 23-24 2017
- Pierre Dragicevic and Yvonne Jansen: the Curated List of Physical Visualizations is continuously being updated.
- Pierre Dragicevic and Jean-Daniel Fekete: Demo of Zooids at Inria's 50th anniversary. 7 Nov 2017.
- Jean-Daniel Fekete and Pierre Dragicevic interviewed by France Culture (la méthode scientifique) on Zooids. 21 Dec 2017.
- The Zooids video has been watched 12 million times, shared 65,000 times and liked 86,000 times on Facebook.

9. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] S. K. BADAM, N. ELMQVIST, J.-D. FEKETE.*Steering the Craft: UI Elements and Visualizations for Supporting Progressive Visual Analytics*, in "Computer Graphics Forum", June 2017, vol. 36, n^o^ 3, p. 491–502 [*DOI :* 10.1111/CGF.13205], https://hal.inria.fr/hal-01512256.

[2] M. BEHRISCH, B. BACH, M. HUND, M. DELZ, L. VON RÜDEN, J.-D. FEKETE, T. SCHRECK.*Magnostics: Image-based Search of Interesting Matrix Views for Guided Network Exploration*, in "IEEE Transactions on Visualization and Computer Graphics", 2017, 1 [*DOI :* 10.1109/TVCG.2016.2598467], https://hal.inria.fr/hal-01377861.

[3] L. BESANÇON, P. ISSARTEL, M. AMMI, T. ISENBERG.*Hybrid Tactile/Tangible Interaction for 3D Data Exploration*, in "IEEE Transactions on Visualization and Computer Graphics", January 2017, vol. 23, n^o^ 1, p. 881–890 [*DOI :* 10.1109/TVCG.2016.2599217], https://hal.inria.fr/hal-01372922.

[4] N. BOUKHELIFA, A. BEZERIANOS, W. CANCINO, E. LUTTON.*Evolutionary Visual Exploration: Evaluation of an IEC Framework for Guided Visual Search*, in "Evolutionary Computation", 2017, to appear [*DOI :* 10.1162/EVCO_A_00161], https://hal.inria.fr/hal-01218959.

[5] L. CRISSAFF, L. WOOD RUBY, S. DEUTCH, R. L. DUBOIS, J.-D. FEKETE, J. FREIRE, C. SILVA.*ARIES: Enabling Visual Exploration and Organization of Art Image Collections*, in "IEEE Computer Graphics and Applications", 2017, 12 [*DOI :* 10.1109/MCG.2017.377152546], https://hal.inria.fr/hal-01612775.

[6] E. DIMARA, A. BEZERIANOS, P. DRAGICEVIC.*Conceptual and Methodological Issues in Evaluating Multidimensional Visualizations for Decision Support*, in "IEEE Transactions on Visualization and Computer Graphics", 2018, vol. 24 [*DOI :* 10.1109/TVCG.2017.2745138], https://hal.inria.fr/hal-01584729.

[7] P. DRAGICEVIC, Y. JANSEN.*Blinded with Science or Informed by Charts? A Replication Study*, in "IEEE Transactions on Visualization and Computer Graphics", January 2018, vol. 24, n^o^ 1, p. 781-790 [*DOI :* 10.1109/TVCG.2017.2744298], http://hal.upmc.fr/hal-01580259.

[8] J. FUCHS, P. ISENBERG, A. BEZERIANOS, D. KEIM.*A Systematic Review of Experimental Studies on Data Glyphs*, in "IEEE Transactions on Visualization and Computer Graphics", July 2017, vol. 23, n^o^ 7, p. 1863-1879 [*DOI :* 10.1109/TVCG.2016.2549018], https://hal.inria.fr/hal-01378429.

[9] P. GOFFIN, J. BOY, W. WILLETT, P. ISENBERG.*An Exploratory Study of Word-Scale Graphics in Data-Rich Text Documents*, in "IEEE Transactions on Visualization and Computer Graphics", 2017, vol. 23, n^o^ 10, p. 2275-2287 [*DOI :* 10.1109/TVCG.2016.2618797], https://hal.inria.fr/hal-01389998.

[10] P. ISENBERG, F. HEIMERL, S. KOCH, T. ISENBERG, P. XU, C. D. STOLPER, M. SEDLMAIR, J. CHEN, T. MÖLLER, J. STASKO.*vispubdata.org: A Metadata Collection about IEEE Visualization (VIS) Publications*, in "IEEE Transactions on Visualization and Computer Graphics", September 2017, vol. 23, n^o^ 9, p. 2199-2206 [*DOI :* 10.1109/TVCG.2016.2615308], https://hal.inria.fr/hal-01376597.

[11] P. ISENBERG, T. ISENBERG, M. SEDLMAIR, J. CHEN, T. MÖLLER.*Visualization as Seen Through its Research Paper Keywords*, in "IEEE Transactions on Visualization and Computer Graphics", January 2017, vol. 23, n^{o} 1, p. 771–780 [*DOI :* 10.1109/TVCG.2016.2598827], https://hal.inria.fr/hal-01370750.

[12] P. ISSARTEL, F. GUÉNIAT, T. ISENBERG, M. AMMI.*Analysis of Locally Coupled 3D Manipulation Mappings Based on Mobile Device Motion*, in "MIT Presence", October 2017, vol. 26, n^{o} 1, p. 66-95 [*DOI :* 10.1162/PRES A 00287], https://hal.inria.fr/hal-01581185.

[13] K. LAWONN, I. VIOLA, T. ISENBERG, B. PREIM.*A Survey of Surface-Based Illustrative Rendering for Visualization*, in "Computer Graphics Forum", 2018, vol. 37, https://hal.inria.fr/hal-01652236.

[14] D. MARTÍN, G. ARROYO, A. RODRÍGUEZ, T. ISENBERG.*A Survey of Digital Stippling*, in "Computers and Graphics", October 2017, vol. 67, p. 24–44 [*DOI :* 10.1016/J.CAG.2017.05.001], https://hal.inria.fr/hal-01528484.

[15] M. MIAO, E. DE LLANO, J. SORGER, Y. AHMADI, T. KEKIC, T. ISENBERG, M. E. GRÖLLER, I. BARIŠIC, I. VIOLA.*Multiscale Visualization and Scale-Adaptive Modification of DNA Nanostructures*, in "IEEE Transactions on Visualization and Computer Graphics", January 2018, vol. 24, n^{o} 1, p. 1014–1024 [*DOI :* 10.1109/TVCG.2017.2743981], https://hal.inria.fr/hal-01581203.

[16] I. VIOLA, T. ISENBERG.*Pondering the Concept of Abstraction in (Illustrative) Visualization*, in "IEEE Transactions on Visualization and Computer Graphics", 2018, vol. 24, forthcoming [*DOI :* 10.1109/TVCG.2017.2747545], https://hal.inria.fr/hal-01581177.

[17] W. WILLETT, Y. JANSEN, P. DRAGICEVIC.*Embedded Data Representations*, in "IEEE Transactions on Visualization and Computer Graphics", 2017, p. 461 - 470 [*DOI :* 10.1109/TVCG.2016.2598608], https://hal.inria.fr/hal-01377901.

[18] E. ZGRAGGEN, A. GALAKATOS, A. CROTTY, J.-D. FEKETE, T. KRASKA.*How Progressive Visualizations Affect Exploratory Analysis*, in "IEEE Transactions on Visualization and Computer Graphics", 2017, vol. 23, n^{o} 8, p. 1977-1987 [*DOI :* 10.1109/TVCG.2016.2607714], https://hal.inria.fr/hal-01377896.

International Conferences with Proceedings

[19] L. BESANÇON, M. AMMI, T. ISENBERG.*Pressure-Based Gain Factor Control for Mobile 3D Interaction using Locally-Coupled Devices*, in "CHI 2017 - ACM CHI Conference on Human Factors in Computing Systems", Denver, United States, May 2017, p. 1831–1842 [*DOI :* 10.1145/3025453.3025890], https://hal.inria.fr/hal-01436172.

[20] L. BESANÇON, P. DRAGICEVIC.*The Significant Difference between p-values and Confidence Intervals*, in "29ème conférence francophone sur l'Interaction Homme-Machine", Poitiers, France, AFIHM (editor), AFIHM, August 2017, 10, Alt.IHM, https://hal.inria.fr/hal-01562281.

[21] L. BESANÇON, P. ISSARTEL, M. AMMI, T. ISENBERG.*Interactive 3D Data Exploration Using Hybrid Tactile/Tangible Input*, in "Journées Visu 2017", Rueil-Malmaison, France, June 2017, https://hal.inria.fr/hal-01529648.

[22] L. BESANÇON, P. ISSARTEL, M. AMMI, T. ISENBERG.*Mouse, Tactile, and Tangible Input for 3D Manipulation*, in "Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)", Den-

ver, United States, May 2017, p. 4727–4740 [*DOI :* 10.1145/3025453.3025863], https://hal.inria.fr/hal-01436206.

[23] E. DIMARA, A. BEZERIANOS, P. DRAGICEVIC.*Narratives in Crowdsourced Evaluation of Visualizations: A Double-Edged Sword?*, in "ACM Conference on Human Factors in Computing Systems (CHI)", Denver, United States, May 2017 [*DOI :* 10.1145/3025453.3025870], https://hal.inria.fr/hal-01448095.

[24] J. JO, J. SEO, J.-D. FEKETE.*A Progressive k-d tree for Approximate k-Nearest Neighbors*, in "Workshop on Data Systems for Interactive Analysis (DSIA)", Phoenix, United States, October 2017, https://hal.inria.fr/hal-01650272.

[25] A. SEMMO, T. ISENBERG, J. DÖLLNER.*Neural Style Transfer: A Paradigm Shift for Image-based Artistic Rendering?*, in "NPAR 2017 - Proceedings of the International Symposium on Non-Photorealistic Animation and Rendering (NPAR as part of Expressive)", Los Angeles, United States, ACM, July 2017, p. 5:1–5:13 [*DOI :* 10.1145/3092919.3092920], https://hal.inria.fr/hal-01527495.

[26] M. SERENO, M. AMMI, T. ISENBERG, L. BESANÇON.*Hybrid Tactile/Tangible Interaction for 3D Selection*, in "29ème conférence francophone sur l'Interaction Homme-Machine", Poitiers, France, AFIHM (editor), AFIHM, August 2017, 6, TEC - Travaux en Cours [*DOI :* 10.1145/3132129.3132150], https://hal.archives-ouvertes.fr/hal-01577940.

Conferences without Proceedings

[27] V. SERRANO MOLINERO, B. BACH, C. PLAISANT, N. DUFOURNAUD, J.-D. FEKETE.*Understanding the Use of The Vistorian: Complementing Logs with Context Mini-Questionnaires*, in "Visualization for the Digital Humanities", Phoenix, United States, October 2017, https://hal.inria.fr/hal-01650259.

Other Publications

[28] E. DIMARA, P. VALDIVIA, C. KINKELDEY.*DCPAIRS: A Pairs Plot Based Decision Support System*, June 2017, EuroVis - 19th EG/VGTC Conference on Visualization, Poster, https://hal.inria.fr/hal-01516470.

[29] N. HEULOT, J.-D. FEKETE, M. AUPETIT.*Visualizing Dimensionality Reduction Artifacts: An Evaluation*, May 2017, https://arxiv.org/abs/1705.05283 - working paper or preprint, https://hal.inria.fr/hal-01523028.

[30] P. ISENBERG, C. KINKELDEY, J.-D. FEKETE.*Exploring Entity Behavior on the Bitcoin Blockchain*, October 2017, p. 1-2, VIS 2017 - IEEE Conference on Visualization, Poster, https://hal.inria.fr/hal-01658500.

[31] C. KINKELDEY, J.-D. FEKETE, P. ISENBERG.*BitConduite: Visualizing and Analyzing Activity on the Bitcoin Network*, June 2017, 3, EuroVis 2017 - Eurographics Conference on Visualization, Posters Track, Poster, https://hal.inria.fr/hal-01528605.

[32] M. SERENO, M. AMMI, T. ISENBERG, L. BESANÇON.*Interaction Hybride Tactile/Tangible pour la Sélection 3D*, August 2017, IHM 2017, Poster, https://hal.inria.fr/hal-01570443.

[33] P. R. VALDIVIA, P. BUONO, J.-D. FEKETE.*Hypenet: Visualizing Dynamic Hypergraphs*, June 2017, p. 1-3, EuroVis 2017 - 19th EG/VGC Conference on Visualization, Poster, https://hal.inria.fr/hal-01658087.

[34] X. WANG, L. BESANÇON, M. AMMI, T. ISENBERG. *Augmenting Tactile 3D Data Exploration With Pressure Sensing*, October 2017, IEEE VIS 2017, Poster, https://hal.inria.fr/hal-01570442.

References in notes

[35] S. K. CARD, J. D. MACKINLAY, B. SHNEIDERMAN (editors). *Readings in information visualization: using vision to think*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1999.

[36] J. BERTIN. *Sémiologie graphique : Les diagrammes - Les réseaux - Les cartes*, Les réimpressions, Editions de l'Ecole des Hautes Etudes en Sciences, Paris, France, 1967.

[37] A. BEZERIANOS, P. DRAGICEVIC, R. BALAKRISHNAN. *Mnemonic rendering: an image-based approach for exposing hidden changes in dynamic displays*, in "UIST '06: Proceedings of the 19th annual ACM symposium on User interface software and technology", New York, NY, USA, ACM, 2006, p. 159–168, http://doi.acm.org/10.1145/1166253.1166279.

[38] J.-D. FEKETE, C. PLAISANT. *Interactive Information Visualization of a Million Items*, in "Proc. IEEE Symposium on Information Visualization 2002 (InfoVis 2002)", Boston, USA, IEEE Press, October 2002, p. 117-124.

[39] J.-D. FEKETE, C. PLAISANT. *Les leçons tirées des deux compétitions de visualisation d'information*, in "Proceedings of IHM 2004", Namur, Belgium, International Conference Proceedings Series, ACM Press, September 2004, p. 7-12.

[40] J. J. GIBSON. *The Ecological Approach to Visual Perception*, Lawrence Erlbaum Associates, New Jersey, USA, 1979.

[41] Y. GUIARD, Y. DU, J.-D. FEKETE, M. BEAUDOUIN-LAFON, C. APPERT, O. CHAPUIS. *Shakespeare's Complete Works as a Benchmark for Evaluating Multiscale Document-Navigation Techniques*, in "Proceedings of BEyond time and errors: novel evaLuation methods for Information Visualization (BELIV'06)", Venice, Italy, ACM Press, May 2006, p. 65-70.

[42] N. HENRY RICHE, J.-D. FEKETE. *Evaluating Visual Table Data Understanding*, in "Proceedings of BEyond time and errors: novel evaLuation methods for Information Visualization (BELIV'06)", Venice, Italy, ACM Press, May 2006, 6.

[43] B. LEE, C. PLAISANT, C. SIMS PARR, J.-D. FEKETE, N. HENRY RICHE. *Task taxonomy for graph visualization*, in "BELIV '06: Proceedings of the 2006 AVI workshop on BEyond time and errors", New York, NY, USA, ACM, 2006, p. 1–5, http://doi.acm.org/10.1145/1168149.1168168.

[44] C. PLAISANT, J.-D. FEKETE, G. GRINSTEIN. *Promoting Insight-Based Evaluation of Visualizations: From Contest to Benchmark Repository*, in "IEEE Transactions on Visualization and Computer Graphics", 2008, vol. 14, n^o^ 1, p. 120–134, http://doi.ieeecomputersociety.org/10.1109/TVCG.2007.70412.

[45] A. TRIESMAN. *Preattentive Processing in Vision*, in "Computer Vision, Graphics, and Image Processing", August 1985, vol. 31, n^o^ 2, p. 156-177.

[46] E. TUFTE. *The Visual Display of Quantitative Information*, Graphics Press, 1983.

[47] J. W. TUKEY.*Exploratory Data Analysis*, Addison-Wesley, 1977.

Team CEDAR

Rich Data Exploration at Cloud Scale

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions).

RESEARCH CENTER
Saclay - Île-de-France

THEME
Data and Knowledge Representation and Processing

Table of contents

Team CEDAR

Creation of the Team: 2016 January 01

Keywords:

Computer Science and Digital Science:

A3.1.1. - Modeling, representation
A3.1.2. - Data management, quering and storage
A3.1.3. - Distributed data
A3.1.6. - Query optimization
A3.1.7. - Open data
A3.1.8. - Big data (production, storage, transfer)
A3.1.9. - Database
A3.2.1. - Knowledge bases
A3.2.3. - Inference
A3.2.4. - Semantic Web
A3.2.5. - Ontologies
A3.3.1. - On-line analytical processing
A3.3.2. - Data mining
A3.3.3. - Big data analysis
A3.4.1. - Supervised learning
A3.4.6. - Neural networks
A3.4.8. - Deep learning
A9.1. - Knowledge
A9.2. - Machine learning

Other Research Topics and Application Domains:

B1.1.6. - Genomics
B8.5.1. - Participative democracy
B9.4.5. - Data science
B9.7.2. - Open data

1. Personnel

Research Scientists

Ioana Manolescu [Team leader, Inria, Senior Researcher, HDR]
Yanlei Diao [Ecole Polytechnique, Senior Researcher, HDR]
Michael Thomazo [Inria, Researcher]

External Collaborators

Ahmed Abdelkafi [Ecole Polytechnique]
Lars Kegel [University of Dresden, from Sep 2017]
Swen Ribeiro [CNRS, until May 2017]
Xavier Tannier [CNRS]
Stamatios Zampetakis [Orchestra Networks, from Sep 2017]

Technical Staff

Oscar Santiago Mendoza Rivera [Inria, until Aug 2017]
Tayeb Merabti [Inria, from Sep 2017]

PhD Students

Maxime Buron [Inria, from Oct 2017]
Tien Duc Cao [Inria]
Sejla Cebiric [Inria]
Luciano Di Palma [Ecole Polytechnique, from Sep 2017]
Felix Raimundo [Ecole Polytechnique, from Sep 2017]
Alexandre Sevin [Ecole Polytechnique, from Oct 2017]
Khaled Zaouk [Ecole Polytechnique, from Oct 2017]
Enhui Huang [Ecole Polytechnique]

Post-Doctoral Fellow

Fei Song [Inria]

Administrative Assistant

Maeva Jeannot [Inria]

2. Overall Objectives

2.1. Overall Objectives

Our research aims at **models, algorithms and tools for highly efficient, easy-to-use data and knowledge management**; throughout our research, **performance at scale** is a core concern, which we address, among other techniques, by designing algorithms for a **cloud (massively parallel)** setting. Our scientific contributions fall in three interconnected areas:

Expressive models for new applications As data and knowledge applications keep extending to novel application areas, we work to devise appropriate data and knowledge models, endowed with formal semantics, to capture such applications' needs. This work mostly concerns the domains of data journalism and journalistic fact checking;

Optimization and performance at scale This topic is at the heart of Y. Diao's ERC project "Big and Fast Data", which aims at optimization with performance guarantees for real-time data processing in the cloud. Machine learning techniques and multi-objectives optimization are leveraged to build performance models for data analytics the cloud. The same boal is shared by our work on efficient evaluation of queries in dynamic knowledge bases.

Data discovery and exploration Today's Big Data is complex; understanding and exploiting it is difficult. To help users, we explore: compact summaries of knowledge bases to abstrac their structure and help users formulate queries; interactive exploration of large relational databases; techniques for automatically discovering interesting information in knowledge bases; and keyword search techniques over Big Data sources.

3. Research Program

3.1. Scalable Heterogeneous Stores

Big Data applications increasingly involve *diverse* data sources, such as: structured or unstructured documents, data graphs, relational databases etc. and it is often impractical to load (consolidate) diverse data sources in a single repository. Instead, interesting data sources need to be exploited "as they are", with the added value of the data being realized especially through the ability to combine (join) together data from several sources. Systems capable of exploiting diverse Big Data in this fashion are usually termed *polystores*. A current limitation of polystores is that data stays captive of its original storage system, which may limit the data exploitation performance. We work to devise highly efficient storage systems for heterogeneous data across a variety of data stores.

3.2. Semantic Query Answering

In the presence of data semantics, query evaluation techniques are insufficient as they only take into account the database, but do not provide the reasoning capabilities required in order to reflect the semantic knowledge. In contrast, (ontology-based) query answering takes into account both the data and the semantic knowledge in order to compute the full query answers, blending query evaluation and semantic reasoning.

We aim at designing efficient semantic query answering algorithms, both building on cost-based reformulation algorithms developed in the team and exploring new approaches mixing materialization and reformulation.

3.3. Multi-Model Querying

As the world's affairs get increasingly more digital, a large and varied set of data sources becomes available: they are either structured databases, such as government-gathered data (demographics, economics, taxes, elections, ...), legal records, stock quotes for specific companies, un-structured or semi-structured, including in particular graph data, sometimes endowed with semantics (see e.g. the Linked Open Data cloud). Modern data management applications, such as data journalism, are eager to combine in innovative ways both static and dynamic information coming from structured, semi-structured, and un-structured databases and social feeds. However, current content management tools for this task are not suited for the task, in particular when they require a lenghy rigid cycle of data integration and consolidation in a warehouse. Thus, we see a need for flexible tools allowing to interconnect various kinds of data sources and to query them together.

3.4. Interactive Data Exploration at Scale

In the Big Data era we are faced with an increasing gap between the fast growth of data and the limited human ability to comprehend data. Consequently, there has been a growing demand of data management tools that can bridge this gap and help users retrieve high-value content from data more effectively. To respond to such user information needs, we aim to build interactive data exploration as a new database service, using an approach called "explore-by-example".

3.5. Exploratory Querying of Semantic Graphs

Semantic graphs including data and knowledge are hard to apprehend for users, due to the complexity of their structure and oftentimes to their large volumes. To help tame this complexity, in prior research (2014), we have presented a full framework for RDF data warehousing, specifically designed for heterogeneous and semantic-rich graphs. However, this framework still leaves to the users the burden of chosing the most interesting warehousing queries to ask. More user-friendly data management tools are needed, which help the user discover the interesting structure and information hidden within RDF graphs.

3.6. Representative Semantic Query Answering

Top-k search is a classical topic, studied in relational databases, semantic web, recommandation systems,... It is extremely useful, among other, when a human user face a large number of query results, allowing the user to reformulate the query if necessary. However, we argue that top-k search incurs a bias on the perception of the set of results which is out of the control of the user. Our goal is to provide the user with k answers as well which are chosen so as to represent the diversity of the answer set. We will first consider this problem in the setting of relational or RDF databases. We will then extend to more heterogeneous sources, including in particular plain text.

4. Application Domains

4.1. Cloud Computing

Cloud computing services are strongly developing and more and more companies and institutions resort to running their computations in the cloud, in order to avoid the hassle of running their own infrastructure. Today's cloud service providers guarantee machine availabilities in their Service Level Agreement (SLA), without any guarantees on performance measures according to a specific cost budget. Running analytics on big data systems require the user not to only reserve the suitable cloud instances over which the big data system will be running, but also setting many system parameters like the degree of parallelism and granularity of scheduling. Chosing values for these parameters, and chosing cloud instances need to meet user objectives regarding latency, throughput and cost measures, which is a complex task if it's done manually by the user. Hence, we need need to transform cloud service models from availabily to user performance objective rises and leads to the problem of multi-objective optimization. Research carried out in the team within the ERC project "Big and Fast Data Analytics" aims to develop a novel optimization framework for providing guarantees on the performance while controlling the cost of data processing in the cloud.

4.2. Computational Journalism

Modern journalism increasingly relies on content management technologies in order to represent, store, and query source data and media objects themselves. Writing news articles increasingly requires consulting several sources, interpreting their findings in context, and crossing links between related sources of information. CEDARresearch results directly applicable to this area provide techniques and tools for rich Web content warehouse management. Within the ANR ContentCheck project, and also as part of our international collaboration with the AIST institute from Japan, we work on one hand, to lay down foundations for computational data journalism and fact checking, and also work to devise concrete algorithms and platforms to help journalists perform their work better and/or faster. This work is carried in collaboration with Le Monde's "Les Décodeurs".

On a related topic, heterogeneous data integration under a virtual graph abstract model is studied within the ICODA Inria project which has started in September 2017. There, we collaborate with Les Décodeurs as well as with Ouest France and Agence France Presse (AFP). The data and knowledge integration framework resulting from this work will support journalists' effort to organize and analyze their knowledge and exploit it in order to produce new content.

4.3. Open Data Intelligence

The Web is a vast source of information, to which more is added every day either in unstructured form (Web pages) or, increasingly, as partially structured sources of information, in particular as Open Data sets, which can be seen as connected graphs of data, most frequently described in the RDF data format recommended by the W3C. Further, RDF data is also the most appropriate format for representing structured information extracted automatically from Web pages, such as the DBPedia database extracted from Wikipedia or Google's InfoBoxes. We work on this topic within the 4-year project ODIN started in 2014.

4.4. Genomics

One particular case of area where the increase in data production is the more consequent is genomic data, indeed the amount of data produced doubles every 7 months. Thus we want to bring the expertise from the database and big data community to help both scale the existing algorithms and design new algorithms that are scalable from the ground up.

5. Highlights of the Year

5.1. Program Committee Chair

Yanlei Diao has been the PC chair of the IEEE International Conference on Data Engineering (ICDE) 2017.

5.2. Strong recruitment of PhD students

The team has started work on many new projects, particularly ; six new PhD thesis starting this year (M. Buron, L. Di Palma, L. Duroyon, F. Raimundo, A. Sevin and K. Zaouk) have rejoined the three more senior students (D. Cao, S. Cebiric, E. Huang). These recruitments boost our efforts on core topics of the team, namely: data exploration, fact checking and data journalism, and performance optimization in the cloud.

5.3. Keynotes

Y. Diao gave a distinguished talk at TU Darmstadt; I. Manolescu gave two keynotes at the international conferences DEXA 2017 and iiWAS 2017.

6. New Software and Platforms

6.1. RDF-Commons

KEYWORDS: Data management - RDF
FUNCTIONAL DESCRIPTION: RDF-Commons is a set of modules providing the abilities to: - load and store RDF data in a DBMS - parse RDF conjunctive queries - encode URIs and literals into integers - encode RDF conjunctive queries - build statistics on RDF data - estimate the cost of the evaluation of a conjunctive query - saturate the RDF data, with respect to an RDF Schema - reformulate a conjunctive query with respect to an RDF Schema - propose algebraic plans

- Contact: Ioana Manolescu

6.2. RDFSummary

FUNCTIONAL DESCRIPTION: RDF Summary is a standalone Java software capable of building summaries of RDF graphs. Summaries are compact graphs (typically several orders of magnitude smaller than the original graph), which can be used to get acquainted quickly with a given graph, they can also be used to perform static query analysis, infer certain things about the answer of a query on a graph, just by considering the query and the summary.

- Contact: Sejla Cebiric

6.3. Tatooine

KEYWORDS: Data integration - Databases - Knowledge database - JSon - RDF - Polystore
FUNCTIONAL DESCRIPTION: Tatooine allows to jointly query data sources of heterogeneous formats and data models (relations, RDF graphs, JSON documents etc.) under a single interface. It is capable of evaluating conjunctive queries over several such data sources, distributing computations between the underlying single-data model systems and a Java-based integration layer based on nested tuples.

- Participants: François Goasdoué, Ioana Manolescu, Javier Letelier Ruiz, Michaël Thomazo, Oscar Santiago Mendoza Rivera, Raphael Bonaque, Swen Ribeiro, Tien Duc Cao and Xavier Tannier
- Contact: Ioana Manolescu

7. New Results

7.1. Semantic Query Answering

Building upon last year's work on regular path queries, we studied the complexity of answering conjunctive regular path queries under linear existential rules and under guarded existential rules. These queries generalized conjunctive queries by their ability to check for a path between two individuals which is labeled by a word belonging to a given regular language. Linear and guarded rules are widely recognized as two important classes of existential rules, that among other generalizes most popular Horn description logics. The results are quite positive, in the sense that the complexity is as good as we could hope for: we provided matching upper-bound that correspond to much less expressive query or ontology languages (i.e., they come from RPQs over linear rules or CQs over guarded rules). These results have been published at IJCAI'17 [13].

7.2. Representative Semantic Query Answers

The availability of large knowledge bases such as Yago or DBPedia allows theoretically anybody to tap in their resources through structured and semantics queries. Thi is still not as widespread as it could be, and we postulate this is mainly for two reasons. First, it is complex to write queries in such a setting. Second, the value added of such querying is improvable. We focused on the second point, with the rationale that increasing the value aded may motivate more easily users to spend the time and energy necessary to learn to write SPARQL queries. More specifically, the internship of M. Buron [22] explored the possibility of exploiting the reasoning performed to find a tuple as an answer to cluster answers in a semantic (and explainable) way.

7.3. Interactive Data Exploration at Scale

To respond to increasing user information needs in the era of Big Data, we aim to build interactive data exploration as a new database service, using an approach called "explore-by-example". In particular, we cast the "explore-by-example" problem in a principled "active learning" framework, and bring the properties of important classes of database queries to bear on the design of new algorithms and optimizations for active learning based database exploration. We introduce a dual-space (data and version space) model for convex pattern queries, leverage the factorized dual-space model and online feature selection to handle high dimensional exploration, and design a new active learning algorithm based on version space reduction. These new techniques allow the database system to not only gain improved accuracy but also overcome fundamental limitations of traditional active learning, in particular, the slow convergence problem. Evaluation results using real-world datasets and user interest patterns show that our new system significantly outperforms state-of-the-art active learning techniques and data exploration systems in accuracy while achieving desired efficiency for interactive performance. In addition, we will extend current data exploration system to handle more complex inputs, such as pictures, by adding a active representation learning phase via neural networks to the existing system. Part of this work was explored during the M2 internship of Alexandre Sevin [25].

7.4. A Quotient Framework for Summarizing RDF Graphs

RDF is the data model of choice for Semantic Web applications. RDF graphs are often large and heterogeneous, thus users may have a hard time determining whether a graph is useful for a certain application. We consider answering such questions by inspecting a *graph summary*, a compact structure conveying as much information as possible about the input graph. A summary is *representative* of a graph if it represents both its explicit and implicit triples, the latter resulting from RDF Schema constraints. To ensure representativeness, we defined a novel RDF-specific summarization framework based on *RDF node equivalence* and graph *quotients*; our framework can be instantiated with many different RDF node equivalence relations. We have shown that our summaries are representative, and establish a *sufficient condition* on the RDF equivalence relation to ensure that a graph can be *efficiently summarized*, without materializing its implicit triples. We illustrate our framework on *bisimulation* equivalence relations between graph nodes, and demonstrate the performance benefits of our efficient summarization method through a set of experiments. These results appeared in [17] and are extended in [20], [19].

7.5. Exploring RDF Graphs through Aggregation

RDF graphs may be large and their structure is heterogeneous and complex, making them very hard to explore and understand. To help users discover valuable insights from RDF graph, we have developed Dagger, a tool which automatically recommends *interesting aggregation queries* over the RDF graphs; Dagger evaluates the queries and graphically shows their results to the user, in the ranked order of their interestingness. To specify aggregate RDF queries, we rely on a dialect of SPARQL 1.1, the standard Semantic Web query language, which has been recently enhanced with the capability to specify aggregation; for the interestingness measure, we relied on variance (or second statistic moment). Dagger was developed as part of the M2 internship of Shu Shang [26] and was demonstrated at the International Semantic Web Conference [15]. A short video of our demo appears online at: https://team.inria.fr/cedar/projects/dagger.

7.6. Models and Algorithms for Fact-Checking and Data Journalism

We have advanced toward a generic definition of a computational fact-checking platform, and identified the set of core functionalities it should support: (i) extraction of a claim from a larger document (typically a text published online in some media, social network etc.); this may require identifying the time and space context in which the claim is supposed to hold; (ii) checking the accuracy of the claim against a set of reference data sources; (iii) putting the claim into perspective by checking its significance in a broader context, for instance by checking if the claim still holds after some minor modification of its temporal, spatial or numeric parameters. Checking a claim is not possible in the absence of a set of reference sources, containing data we consider to be true; thus reference source construction, refinement and selection are also central tasks in such an architecture. We have carried this work as part of the ANR ContentCheck project (Section 8.1.1) and also within our associated team with AIST Japan (Section 8.2.1.1). The architecture of the generic platform we envision has been presented in the Paris DB Day event in May 2017, in an ERCIM News [21] and in a keynote [24].

Within this architecture, an important task is to construct reference data sources and to make them more accessible. Toward this goal, we have devised an approach to extract Linked Open Data (RDF graphs) from Excel tables published by INSEE, the French national statistics institute [14]; the resulting data has been published online. Another ongoing line of work explored within the PhD of Ludivine Duroyon concerns establishing new models for temporal beliefs and statements, allowing journalists to increase the value of refernece sources on which to check who said what when.

7.7. Design and optimization for population genomics

As mentioned above, the area of genomics experiences a massive increase in the amount of data to be processed. Furthermore the data generated can sometimes hard to interpret (in particular NGS data for CNV detection).

We investigate new means to discover Copy Number Variation in the human population using methods from the deep learning community. Indeed, great success has been achieved in that area within projects such as DeepVariant; such projects managed to considerably lower the latency for getting results (about 10 fold) but at a higher computational cost. Such methods are currently attracting significant attention in the biology / bioinformatics community, as witnesed by an editorial in Cell Systems (December 2017) [0].

As the area of population genomics is fairly new, we hope to help design a complete framework allowing for better optimisations and integration with database tools. This work is carried by Yanlei Diao and Felix Raimundo, together with Dr. Avinash Abhyankar at the New York Genome Center (NYGC) who co-advises the PhD of F. Raimundo and Dr. Toby Bloom (head of informatics at NYGC).

7.8. Performance Modeling and Multi-Objective Optimization For the Cloud

We study cloud service models based on attaining user's performance objectives; these immediately lead to problems of multi-objective optimization.

[0] http://www.cell.com/cell-systems/fulltext/S2405-4712(17)30554-9

Given different cost models, we consider the optimizer will search a multi-dimensional space, compute execution plans that are not dominated by others (known as Pareto plans) and explore meaningful tradeoffs between different objectives to find the optimal plan for each analytical task. We focused on analytical tasks encoded as dataflow programs as in Hadoop and Spark systems. When such dataflow programs are submitted to the cloud, we aim to provide a multi-objective optimizer that can automatically find an optimal execution plan of the dataflow program, which meets specific user performance objectives. Developing an optimizer for dataflow programs in the cloud raises two major challenges: The optimizer needs cost models for running complex dataflow programs in the cloud, and, it further needs a new algorithmic foundation for multi-objective optimization across user-specific objectives.

We have worked to develop a performance model for the optimizer in order to build the skylines for the user-objectives. We found that deep learning offers an incremental prediction framework (using embedding architecture) or online prediction framework (using auto-encoder along with a gradient boosting regressor) that are not available in a baseline regressor approach. However, there is a tradeoff between using the online prediction framework and having good performance, since of course retraining improves results. That said, the online prediction framework gave us acceptable generalization power over unseen jobs. This work has been carried in the M2 internship of Khaled Zaouk [27], and it continues through his PhD.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

- AIDE ("A New Database Service for Interactive Exploration on Big Data") is an ANR "Young Researcher" project led by Y. Diao, started at the end of 2016.
- CBOD ("Cloud-Based Organizational Design") is a 4-year ANR started in 2014, coordinated by prof. Ahmed Bounfour from UPS. Its goal is to study and model the ways in which cloud computing impacts the behavior and operation of companies and organizations, with a particular focus on the cloud-based management of data, a crucial asset in many companies.
- ContentCheck (2015-2018) is an ANR project in collaboration with U. Rennes 1 (F. Goasdoué), INSA Lyon (P. Lamarre), the LIMSI lab from U. Paris Sud, and the Le Monde newspaper, in particular their fact-checking team Les Décodeurs. Its aim is to investigate content management models and tools for journalistic fact-checking.

8.1.2. LabEx, IdEx

- CloudSelect is a three-years project started in October 2015. It is financed by the *Institut de la Société Numérique* (ISN) of the IDEX Paris-Saclay; it funds the PhD scholarship of S. Cebiric. The project is a collaboration with A. Bounfour from the economics department of Université Paris Sud. The project aims at exploring technical and business-oriented aspects of data mobility across clo ud services, and from the cloud to outside the cloud.

8.1.3. Others

- ODIN is a four-year project started in 2014, funded by the Direction Générale de l'Armement, between the SemSoft company, IRISA Rennes and Cedar. The project aims to develop a complete framework for analytics on Web data, in particular taking into account uncertainty, based on Semantic Web technologies such as RDF.
- The goal of the iCODA project is to develop the scientific and technological foundations for knowledge- mediated user-in-the-loop collaborative data analytics on heterogenous information sources, and to demonstrate the effectiveness of the approach in realistic, high-visibility use-cases. The project stands at the crossroad of multiple research fields—content analysis, data management,

knowledge represen- tation, visualization—that span multiple Inria themes, and counts on a club of major press partners to define usage scenarios, provide data and demonstrate achievements. This is a project funded directly by Inria ("Inria Project Lab"), and is in collaboration with GraphIK, ILDA, LINKMEDIA (coordinator), as well as the press partners AFP, Le Monde (Les Décodeurs) and Ouest-France.

8.2. International Initiatives

8.2.1. Inria Associate Teams Not Involved in an Inria International Labs

8.2.1.1. WebClaimExplain

Title: Mining for explanations to claims published on the Web

International Partner (Institution - Laboratory - Researcher):

AIST (Japan) - Julien Leblay

Start year: 2017

See also: https://team.inria.fr/cedar/projects/webclaimexplain/

The goal of this research is to create tools to find explanations for facts and verify claims made online. While this process cannot be fully automated, the main focus of our work will be explanation finding via trusted sources, based on the observation that one can only trust a statement if he/she can explain it through rules and proofs that can themselves be trusted.

8.3. International Research Visitors

8.3.1. Visits of International Scientists

8.3.1.1. Internships

Lars Kegel, a PhD student at the university of Dresden, is visiting the team since September 2017. He is working on the systematic description of time series with features that capture the global, structural characteristics of a series in a lower dimensional space.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Selection

9.1.1.1. Chair of Conference Program Committees

Y. Diao has been the PC co-chair of IEEE International Conference of Data Engineering (ICDE) 2017.

I. Manolescu has been a track chair for the ACM SIGMOD (Special Interest Group on the Management of Data) 2017 conference.

9.1.1.2. Member of the Conference Program Committees

I. Manolescu has been member of the PC of the DASFAA Conference 2017, of the Semantic Big Data Workshop (SBD) and of the WebDB (Web and Databases) Workshops in conjunction with the SIGMOD Conference 2017.

M. Thomazo has been member of the PC of BDA 2017, and IJCAI 2017.

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards

Y. Diao is the editor-in-chief of ACM SIGMOD Record.

I. Manolescu has been an Associate Editor for PVLDB (Proceedings of Very Large Databases) 2017.

9.1.2.2. Reviewer - Reviewing Activities

M. Thomazo has been a reviewer for JODS and TOCL.

9.1.3. Invited Talks

- Y. Diao gave a distinguished lecture at the Technische Universitaet Darmstadt, in November 2017.
- I. Manolescu gave a keynote talk at the DEXA Conference in August 2017 [23].
- I. Manolescu gave a keynote talk at the iiWAS Conference in December 2017 [24]

9.1.4. Leadership within the Scientific Community

Y. Diao is a member of the ACM SIGMOD Executive Committee, and also a member of the PVLDB Endowment.

I. Manolescu is a member of the PVLDB Endowment, of the ACM SIGMOD "Jim Gray" PhD Award Committee, and of the steering committee (*Comité de Pilotage*) of "Bases de Données Avancées" (BDA), the informal association organizing the database research community in France and french-speaking countries.

9.1.5. Research Administration

Y. Diao is on the advisory board of the Data Science Initiative (DSI), a joint center between the applied mathematics and computer science departments of Ecole Polytechnique.

I. Manolescu is responsible of the "Massive Data Processing" axis of the Inria partnership with DGA (Direction Générale de l'Armement).

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

- Master: Y. Diao is a Professor at Ecole Polytechnique, where she teaches "System for Big Data" in M1; she also teaches "Systems for Big Data Analytics" in M2 in the Data Science Master Program of Université Paris Saclay, and organizes a data science research project of M1 (third-year) students at École Polytechnique.
- Master: I. Manolescu, Architectures for Massive Data Management, 12h, M2, Université Paris-Saclay.
- Master: I. Manolescu, Database Management Systems, 52h, M1, École Polytechnique.

9.2.2. Supervision

PhD in progress: Maxime Buron: "Raisonnement efficace sur des grands graphes hétérogènes ", since October 2017, François Goasdoué, Ioana Manolescu and Marie-Laure Mugnier (GraphIK Inria team in Montpellier)

PhD in progress: Tien Duc Cao: "Extraction et interconnexion de connaissances appliquée aux données journalistiques", since October 2016, Ioana Manolescu and Xavier Tannier (LIMSI/CNRS and Université de Paris Sud)

PhD in progress: Sejla Čebirić: "CloudSelect: Data Mobility Within, Across and Outside Clouds", since September 2015, François Goasdoué Goasdoué and Ioana Manolescu.

PhD in progress: Ludivine Duroyon: "Data management models, algorithms & tools for fact-checking", since October 2017, François Goasdoué and Ioana Manolescu (Ludivine is in the Shaman team of U. Rennes 1 and IRISA, in Lannion)

PhD in progress: Enhui Huang: "Interactive Data Exploration at Scale", since October 2016, Yanlei Diao and Anna Liu (U. Massachussets at Amherst, USA)

PhD in progress: Luciano di Palma, “New sampling algorithms and optimizations for interactive exploration in Big Data”, since October 2017, Yanlei Diao and Anna Liu (U. Massachussets at Amherst, USA)

PhD in progress: Felix Raimundo: “Nouveaux algorithmes et optimisations pour l’analyse profonde du génome à l’échelle de la population”, since October 2017, Yanlei Diao and Avinash Abhyankar (New York Genome Center, USA)

PhD in progress: Alexandre Sevin: “Exploration interactive de données sur de grandes sources de données hétérogènes”, since October 2017, Yanlei Diao and Peter Haas (U. Massachussets at Amherst, USA)

PhD in progress: Khaled Zaouk: “Performance Modeling and Multi-Objective Optimization for Data Analytics in the Cloud”, since October 2017, Yanlei Diao

9.2.3. Juries

Y. Diao has been a member of the PhD committee of Julien Pilourdault (defended in September 2017, at LIG, Grenoble). The thesis was titled “Scalable Algorithms for Monitoring Activity Traces”.

I. Manolescu has been a member of the PhD committee of Olivier Wang (defended in June 2017, at LIX), the thesis was titled “Adaptive Rule Models: Active Learning for Rule-Based Systems” and also of the PhD committee of Maria Rossi (defended in November 2017, at LIX). The thesis was titled “Graph Mining for Influence Maximization in Social Networks”.

I. Manolescu has been a reviewer and a member of the HDR committee of Vicent Leroy (defended in September 2017, at LIG, Grenoble). The thesis was titled “Data Analysis at Scale: Systems, Algorithms and Information”.

9.3. Popularization

- M. Buron, I. Manolescu, F. Raimundo and T. Merabti animated a booth at the “Fête de la Science 2017” at Inria Saclay, in October 2017.
- I. Manolescu published an article on computational fact-checking titled “La vérité, rien que la vérité” in the Binaire blog of Le Monde (http://binaire.blog.lemonde.fr/2017/04/05/la-verite-rien-que-la-verite/)
- I. Manolescu participated to a panel on Data Journalism at the Web2Day, a 3000-strong IT and digital conference in Nantes, in June 2017 (https://web2day.co/en/speakers/ioana-manolescu/)
- I. Manolescu gave a talk at the Inria Alumni Jam Session “ Fausses informations, post vérité : allons aux faits ! ”, presenting our current work on data management for data journalism (http://www.inria-alumni.fr/evenement/session-inria-alumni-fausses-informations-post-verite-allons-aux-faits-25-octobre-2017-cnam-paris/)

10. Bibliography

Major publications by the team in recent years

[1] J.-F. BAGET, M. BIENVENU, M.-L. MUGNIER, M. THOMAZO.*Answering Conjunctive Regular Path Queries over Guarded Existential Rules*, in "IJCAI: International Joint Conference on Artificial Intelligence", Melbourne, Australia, August 2017, https://hal-lirmm.ccsd.cnrs.fr/lirmm-01632224.

[2] R. BONAQUE, B. CAUTIS, F. GOASDOUÉ, I. MANOLESCU.*Toward Social, Structured and Semantic Search*, in "Surfacing the Deep and the Social Web (SDSW)", Riva del Garda, Trentino, Italy, Proceedings of the Workshop on Surfacing the Deep and the Social Web co-located with the 13th International Semantic Web Conference (ISWC 2014), COST Action KEYSTONE, October 2014, vol. 1310, https://hal.inria.fr/hal-01109123.

[3] D. BURSZTYN, F. GOASDOUÉ, I. MANOLESCU.*Teaching an RDBMS about ontological constraints*, in "Proceedings of the Very Large Databases (PVLDB)", 2016, vol. 9, n^{o} 12, p. 1161–1172, http://www.vldb.org/pvldb/vol9/p1161-bursztyn.pdf.

[4] D. COLAZZO, F. GOASDOUÉ, I. MANOLESCU, A. ROATIS.*RDF Analytics: Lenses over Semantic Graphs*, in "23rd International World Wide Web Conference", Seoul, South Korea, April 2014 [*DOI :* 10.1145/2566486.2567982], https://hal.inria.fr/hal-00960609.

[5] Y. DIAO, I. MANOLESCU, S. SHANG.*Dagger: Digging for Interesting Aggregates in RDF Graphs*, in "International Semantic Web Conference (ISWC)", Vienna, Austria, October 2017, https://hal.inria.fr/hal-01577464.

[6] F. GOASDOUÉ, Z. KAOUDI, I. MANOLESCU, J.-A. QUIANÉ-RUIZ, S. ZAMPETAKIS.*CliqueSquare: Flat Plans for Massively Parallel RDF Queries*, in "International Conference on Data Engineering", Seoul, South Korea, April 2015, https://hal.inria.fr/hal-01108705.

[7] K. KARANASOS, A. KATSIFODIMOS, I. MANOLESCU.*Delta: Scalable Data Dissemination under Capacity Constraints*, October 2013, n^{o} RR-8385, 37, https://hal.inria.fr/hal-00877758.

[8] M. KÖNIG, M. LECLÈRE, M. MUGNIER, M. THOMAZO.*Sound, complete and minimal UCQ-rewriting for existential rules*, in "Semantic Web", 2015, vol. 6, n^{o} 5, p. 451–475, http://dx.doi.org/10.3233/SW-140153.

[9] A. ROY, Y. DIAO, U. EVANI, A. ABHYANKAR, C. HOWARTH, R. LE PRIOL, T. BLOOM.*Massively Parallel Processing of Whole Genome Sequence Data: An In-Depth Performance Study*, in "SIGMOD '17 Proceedings of the 2017 ACM International Conference on Management of Dat", Chicago, Illinois, United States, SIGMOD '17 Proceedings of the 2017 ACM International Conference on Management of Data, ACM, May 2017, p. 187-202 [*DOI :* 10.1145/3035918.3064048], https://hal.inria.fr/hal-01683398.

[10] H. ZHANG, Y. DIAO, N. IMMERMAN.*On complexity and optimization of expensive queries in complex event processing*, in "International Conference on Management of Data, SIGMOD 2014, Snowbird, UT, USA, June 22-27, 2014", 2014, p. 217–228, http://doi.acm.org/10.1145/2588555.2593671.

Publications of the year

Articles in International Peer-Reviewed Journal

[11] M. KRÖTZSCH, T. MASOPUST, M. THOMAZO.*Complexity of universality and related problems for partially ordered NFAs*, in "Information and Computation", 2017, vol. 255, p. 177 - 192 [*DOI :* 10.1016/J.IC.2017.06.004], https://hal.inria.fr/hal-01571398.

[12] T. MASOPUST, M. THOMAZO.*On Boolean Combinations forming Piecewise Testable Languages*, in "Theoretical Computer Science", June 2017, vol. 682, https://hal.inria.fr/hal-01637057.

International Conferences with Proceedings

[13] J.-F. BAGET, M. BIENVENU, M.-L. MUGNIER, M. THOMAZO.*Answering Conjunctive Regular Path Queries over Guarded Existential Rules*, in "IJCAI: International Joint Conference on Artificial Intelligence", Melbourne, Australia, August 2017, https://hal-lirmm.ccsd.cnrs.fr/lirmm-01632224.

[14] T. D. CAO, I. MANOLESCU, X. TANNIER.*Extracting linked data from statistic spreadsheets*, in "International Workshop on Semantic Big Data", Chicago, United States, International Workshop on Semantic Big Data, May 2017, p. 1 - 5 [*DOI :* 10.1145/3066911.3066914], https://hal.inria.fr/hal-01583975.

[15] Y. DIAO, I. MANOLESCU, S. SHANG.*Dagger: Digging for Interesting Aggregates in RDF Graphs*, in "International Semantic Web Conference (ISWC)", Vienna, Austria, October 2017, https://hal.inria.fr/hal-01577464.

[16] A. ROY, Y. DIAO, U. EVANI, A. ABHYANKAR, C. HOWARTH, R. LE PRIOL, T. BLOOM.*Massively Parallel Processing of Whole Genome Sequence Data: An In-Depth Performance Study*, in "SIGMOD '17 Proceedings of the 2017 ACM International Conference on Management of Dat", Chicago, Illinois, United States, ACM, May 2017, p. 187-202 [*DOI :* 10.1145/3035918.3064048], https://hal.inria.fr/hal-01683398.

[17] Š. ČEBIRIĆ, F. GOASDOUÉ, I. MANOLESCU.*A Framework for Efficient Representative Summarization of RDF Graphs*, in "International Semantic Web Conference (ISWC)", Vienna, Austria, October 2017, https://hal.inria.fr/hal-01577778.

Research Reports

[18] T. D. CAO, I. MANOLESCU, X. TANNIER.*Extracting Linked Data from statistic spreadsheets*, Inria Saclay Ile de France, March 2017, https://hal.inria.fr/hal-01496700.

[19] Š. ČEBIRIĆ, F. GOASDOUÉ, I. MANOLESCU.*A Framework for Efficient Representative Summarization of RDF Graphs*, Inria Saclay Ile de France ; Ecole Polytechnique ; Université de Rennes 1 [UR1], August 2017, n^{o} RR-9090, 11, https://hal.inria.fr/hal-01577431.

[20] Š. ČEBIRIĆ, F. GOASDOUÉ, I. MANOLESCU.*Query-Oriented Summarization of RDF Graphs*, Inria Saclay ; Université Rennes 1, June 2017, n^{o} RR-8920, https://hal.inria.fr/hal-01325900.

Scientific Popularization

[21] I. MANOLESCU.*ContentCheck: Content Management Techniques and Tools for Fact-checking*, in "ERCIM News", October 2017, https://hal.inria.fr/hal-01596563.

Other Publications

[22] M. BURON.*Grouping Answers in Ontology-Based Query Answering*, Inria Saclay, September 2017, 20, https://hal.inria.fr/hal-01622564.

[23] I. MANOLESCU.*Data Discovery in RDF Graphs*, August 2017, p. 1-63, DEXA 2017 - 28th International Conference on Database and Expert System Applications, https://hal.inria.fr/hal-01657144.

[24] I. MANOLESCU.*Data integration for journalism: goals, tools, and architectures (Keynote)*, December 2017, p. 1-46, iiWAS 2017 - 19th International Conference on Information Integration and Web-based Applications & Services, https://hal.inria.fr/hal-01657152.

[25] A. SEVIN.*Creation of a smart representation of pictures for interactive data exploration*, ENSAE ParisTech, September 2017, https://hal.inria.fr/hal-01643077.

[26] S. SHANG.*Exploratory Analytics for RDF Graphs*, Université de Paris Saclay, September 2017, https://hal.inria.fr/hal-01657163.

[27] K. ZAOUK.*Performance Modeling and Multi-Objective Optimization for Data Analytics in the Cloud*, Ecole Polytechnique (Palaiseau, France) ; Telecom ParisTech, September 2017, https://hal.inria.fr/hal-01647208.

Project-Team COMETE

Concurrency, Mobility and Transactions

IN COLLABORATION WITH: Laboratoire d'informatique de l'école polytechnique (LIX)

IN PARTNERSHIP WITH:
CNRS

Ecole Polytechnique

RESEARCH CENTER
Saclay - Île-de-France

THEME
Security and Confidentiality

Table of contents

Project-Team COMETE

Creation of the Project-Team: 2008 January 01

Keywords:

Computer Science and Digital Science:
A2.1.1. - Semantics of programming languages
A2.1.5. - Constraint programming
A2.1.6. - Concurrent programming
A2.1.8. - Synchronous languages
A2.4.1. - Analysis
A2.4.2. - Model-checking
A3.4. - Machine learning and statistics
A4.1. - Threat analysis
A4.5. - Formal methods for security
A4.8. - Privacy-enhancing technologies

Other Research Topics and Application Domains:
B6.1. - Software industry
B6.6. - Embedded systems
B9.4.1. - Computer science
B9.8. - Privacy

1. Personnel

Research Scientists
Catuscia Palamidessi [Team leader, Inria, Senior Researcher]
Konstantinos Chatzikokolakis [CNRS, Researcher]
Frank Valencia [CNRS, Researcher]

PhD Students
Michell Guzman [Inria]
Anna Pazii [Ecole polytechnique, from Oct 2017]
Tymofii Prokopenko [Inria]
Marco Romanelli [Inria, from Oct 2017]

Post-Doctoral Fellow
Ali Kassem [Inria, from Jun 2017]

Visiting Scientists
Giovanni Cherubin [Royal Holloway University of London, UK, Nov 2017]
Mario Ferreira Alvim Junior [Federal University of Minas Gerais, Brazil, Dec 2017]
David Frutos Escrig [Universidad Complutense Madrid, Spain, until Feb 2017]
Yusuke Kawamoto [AIST, Japan, Nov 2017]
Santiago Quintero [Universidad Javeriana de Cali, Colombia, from Nov to Dec 2017]

Administrative Assistants
Jessica Gameiro [Inria]
Hélèna Kutniak [Inria]

2. Overall Objectives

2.1. Overall Objectives

Our times are characterized by the massive presence of highly *distributed systems* consisting of diverse and specialized devices, forming heterogeneous networks, and providing different services and applications. Revolutionary phenomena such as *social networks* and *cloud computing* are examples of such systems.

In Comète we study emerging concepts of this new era of computing. *Security* and *privacy* are some of the fundamental concerns that arise in this setting. In particular, in the modern digital world the problem of keeping information secret or confidential is exacerbated by orders of magnitude: the frequent interaction between users and electronic devices, and the continuous connection between these devices and the internet, offer malicious agents the opportunity to gather and store huge amount of information, often without the individual even being aware of it. Mobility is an additional source of vulnerability, since tracing may reveal significant information. To avoid these kinds of hazards, *security protocols* and various techniques for privacy protection have been designed. However, the properties that they are supposed to ensure are rather subtle, and, furthermore, it is difficult to foresee all possible expedients that a potential attacker may use. As a consequence, even protocols that seem at first "obviously correct" are later (often years later) found to be prone to attacks.

In addition to the security problems, the problems of correctness, robustness and reliability are made more challenging by the complexity of these systems, since they are highly concurrent and distributed. Despite being based on impressive engineering technologies, they are still prone to faulty behavior due to errors in the software design.

To overcome these drawbacks, we need to develop formalisms, reasoning techniques, and verification methods, to specify systems and protocols, their intended properties, and to guarantee that these intended properties of correctness and security are indeed satisfied.

In Comète we study formal computational frameworks for specifying these systems, theories for defining the desired properties of correctness and security and for reasoning about them, and methods and techniques for proving that a given system satisfies the intended properties.

3. Research Program

3.1. Probability and information theory

Participants: Konstantinos Chatzikokolakis, Catuscia Palamidessi, Marco Romanelli, Anna Pazii.

Much of the research of Comète focuses on security and privacy. In particular, we are interested in the problem of the leakage of secret information through public observables.

Ideally we would like systems to be completely secure, but in practice this goal is often impossible to achieve. Therefore, we need to reason about the amount of information leaked, and the utility that it can have for the adversary, i.e. the probability that the adversary is able to exploit such information.

The recent tendency is to use an information theoretic approach to model the problem and define the leakage in a quantitative way. The idea is to consider the system as an information-theoretic *channel*. The input represents the secret, the output represents the observable, and the correlation between the input and output (*mutual information*) represents the information leakage.

Information theory depends on the notion of entropy as a measure of uncertainty. From the security point of view, this measure corresponds to a particular model of attack and a particular way of estimating the security threat (vulnerability of the secret). Most of the proposals in the literature use Shannon entropy, which is the most established notion of entropy in information theory. We, however, consider also other notions, in particular Rényi min-entropy, which seems to be more appropriate for security in common scenarios like one-try attacks.

3.2. Expressiveness of Concurrent Formalisms

Participants: Catuscia Palamidessi, Frank Valencia.

We study computational models and languages for distributed, probabilistic and mobile systems, with a particular attention to expressiveness issues. We aim at developing criteria to assess the expressive power of a model or formalism in a distributed setting, to compare existing models and formalisms, and to define new ones according to an intended level of expressiveness, also taking into account the issue of (efficient) implementability.

3.3. Concurrent constraint programming

Participants: Michell Guzman, Frank Valencia.

Concurrent constraint programming (ccp) is a well established process calculus for modeling systems where agents interact by posting and asking information in a store, much like in users interact in *social networks*. This information is represented as first-order logic formulae, called constraints, on the shared variables of the system (e.g., $X > 42$). The most distinctive and appealing feature of ccp is perhaps that it unifies in a single formalism the operational view of processes based upon process calculi with a declarative one based upon first-order logic. It also has an elegant denotational semantics that interprets processes as closure operators (over the set of constraints ordered by entailment). In other words, any ccp process can be seen as an idempotent, increasing, and monotonic function from stores to stores. Consequently, ccp processes can be viewed as: computing agents, formulae in the underlying logic, and closure operators. This allows ccp to benefit from the large body of techniques of process calculi, logic and domain theory.

Our research in ccp develops along the following two lines:

1. **(a)** The study of a bisimulation semantics for ccp. The advantage of bisimulation, over other kinds of semantics, is that it can be efficiently verified.
2. **(b)** The extension of ccp with constructs to capture emergent systems such as those in social networks and cloud computing.

3.4. Model checking

Participants: Konstantinos Chatzikokolakis, Catuscia Palamidessi.

Model checking addresses the problem of establishing whether a given specification satisfies a certain property. We are interested in developing model-checking techniques for verifying concurrent systems of the kind explained above. In particular, we focus on security and privacy, i.e., on the problem of proving that a given system satisfies the intended security or privacy properties. Since the properties we are interested in have a probabilistic nature, we use probabilistic automata to model the protocols. A challenging problem is represented by the fact that the interplay between nondeterminism and probability, which in security presents subtleties that cannot be handled with the traditional notion of a scheduler,

4. Application Domains

4.1. Security and privacy

Participants: Konstantinos Chatzikokolakis, Catuscia Palamidessi, Ali Kassem, Anna Pazii, Tymofii Prokopenko.

The aim of our research is the specification and verification of protocols used in mobile distributed systems, in particular security protocols. We are especially interested in protocols for *information hiding*.

Information hiding is a generic term which we use here to refer to the problem of preventing the disclosure of information which is supposed to be secret or confidential. The most prominent research areas which are concerned with this problem are those of *secure information flow* and of *privacy*.

Secure information flow refers to the problem of avoiding the so-called *propagation* of secret data due to their processing. It was initially considered as related to software, and the research focussed on type systems and other kind of static analysis to prevent dangerous operations, Nowadays the setting is more general, and a large part of the research effort is directed towards the investigation of probabilistic scenarios and treaths.

Privacy denotes the issue of preventing certain information to become publicly known. It may refer to the protection of *private data* (credit card number, personal info etc.), of the agent's identity (*anonymity*), of the link between information and user (*unlinkability*), of its activities (*unobservability*), and of its *mobility* (*untraceability)*.

The common denominator of this class of problems is that an adversary can try to infer the private information (*secrets*) from the information that he can access (*observables*). The solution is then to obfuscate the link between secrets and observables as much as possible, and often the use randomization, i.e. the introduction of *noise*, can help to achieve this purpose. The system can then be seen as a *noisy channel*, in the information-theoretic sense, between the secrets and the observables.

We intend to explore the rich set of concepts and techniques in the fields of information theory and hypothesis testing to establish the foundations of quantitive information flow and of privacy, and to develop heuristics and methods to improve mechanisms for the protection of secret information. Our approach will be based on the specification of protocols in the probabilistic asynchronous π-calculus, and the application of model-checking to compute the matrices associated to the corresponding channels.

5. New Software and Platforms

5.1. Location Guard

KEYWORDS: Privacy - Geolocation - Browser Extensions
SCIENTIFIC DESCRIPTION: The purpose of Location Guard is to implement obfuscation techniques for achieving location privacy, in a an easy and intuitive way that makes them available to the general public. Various modern applications, running either on smartphones or on the web, allow third parties to obtain the user's location. A smartphone application can obtain this information from the operating system using a system call, while web application obtain it from the browser using a JavaScript call.
FUNCTIONAL DESCRIPTION: Websites can ask the browser for your location (via JavaScript). When they do so, the browser first asks your permission, and if you accept, it detects your location (typically by transmitting a list of available wifi access points to a geolocation provider such as Google Location Services, or via GPS if available) and gives it to the website.

Location Guard is a browser extension that intercepts this procedure. The permission dialog appears as usual, and you can still choose to deny. If you give permission, then Location Guard obtains your location and adds "random noise" to it, creating a fake location. Only the fake location is then given to the website.

In 2017 there was a major update to the Firefox version of Location Guard, to make it compatible with the Firefox Quantum. This latest Firefox version discontinued support for the legacy addon API, so Location Guard had to be adapted to the new WebExtensions API.

Moreover, the latest version implements new features requested by users, such as the ability to search for a fixed location, as well as bugfixes.

- Participants: Catuscia Palamidessi, Konstantinos Chatzikokolakis, Marco Stronati, Miguel Andrés and Nicolas Bordenabe
- Contact: Konstantinos Chatzikokolakis
- URL: https://github.com/chatziko/location-guard

5.2. libqif - A Quantitative Information Flow C++ Toolkit Library

KEYWORDS: Information leakage - Privacy - C++ - Linear optimization

FUNCTIONAL DESCRIPTION: The goal of libqif is to provide an efficient C++ toolkit implementing a variety of techniques and algorithms from the area of quantitative information flow and differential privacy. We plan to implement all techniques produced by Comète in recent years, as well as several ones produced outside the group, giving the ability to privacy researchers to reproduce our results and compare different techniques in a uniform and efficient framework.

Some of these techniques were previously implemented in an ad-hoc fashion, in small, incompatible with each-other, non-maintained and usually inefficient tools, used only for the purposes of a single paper and then abandoned. We aim at reimplementing those – as well as adding several new ones not previously implemented – in a structured, efficient and maintainable manner, providing a tool of great value for future research. Of particular interest is the ability to easily re-run evaluations, experiments and case-studies from all our papers, which will be of great value for comparing new research results in the future.

The library's development continued in 2017 with several new added features. The project's git repository shows for this year 33 commits by 2 contributors. The new functionality was directly applied to the exeperimental results of several publications of the team (PETS'17, GameSec'17, VALUETOOLS'17).

- Contact: Konstantinos Chatzikokolakis
- URL: https://github.com/chatziko/libqif

5.3. dspacenet

Distributed-Spaces Network.
KEYWORDS: Social networks - Distributed programming
FUNCTIONAL DESCRIPTION: DSpaceNet is a tool for social networking based on multi-agent spatial and timed concurrent constraint language.

I - The fundamental structure of DSPaceNet is that of *space*: A space may contain

(1) spatial-mobile-reactive tcc programs, and (2) other spaces.

Furthermore, (3) each space belongs to a given agent. Thus, a space of an agent j within the space of agent i means that agent i allows agent j to use a computation sub-space within its space.

II - The fundamental operation of DSPaceNet is that of *program posting*: In each time unit, agents can post spatial-mobile-reactive tcc programs in the spaces they are allowed to do so (ordinary message posting corresponds to the posting of tell processes). Thus, an agent can for example post a watchdog tcc process to react to messages in their space, e.g. whenever (*happy b*frank*) do tell("thank you!"). More complex mobile programs are also allowed (see below).

The language of programs is a spatial mobile extension of tcc programs:

$P, Q ... := tell(c) | whencdoP || nextP | P || Q | unlesscnextP | [P]\_i | \uparrow \_iP | recX.P$

computation of timed processes proceeds as in tcc. The spatial construct [P]_i runs P in the space of agent i and the mobile process uparrow_i P, extrudes P from the space of i. By combining space and mobility, arbitrary processes can be moved from one a space into another. For example, one could send a trojan watchdog to another space for spying for a given message and report back to one's space.

III- Constraint systems can be used to specify advance text message deduction, arithmetic deductions, scheduling, etc.

IV - Epistemic Interpretation of spaces can be used to derive whether they are users with conflicting/inconsistent information, or whether a group of agents may be able to deduce certain message.

V - The scheduling of agent requests for program posts, privacy settings, friendship lists are handled by an external interface. For example, one could use type systems to check whether a program complies with privacy settings (for example checking that the a program does not move other program into a space it is not allowed into).

- Partner: Pontificia Universidad Javeriana Cali
- Contact: Frank Valencia
- URL: http://www.dspacenet.com

6. New Results

6.1. Foundations of information hiding

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

6.1.1. Information Leakage Games

In [19] we studied a game-theoretic setting to model the interplay between attacker and defender in the context of information flow, and to reason about their optimal strategies. In contrast with standard game theory, in our games the utility of a mixed strategy is a convex function of the distribution on the defender's pure actions, rather than the expected value of their utilities. Nevertheless, the important properties of game theory, notably the existence of a Nash equilibrium, still hold for our (zero-sum) leakage games, and we provided algorithms to compute the corresponding optimal strategies. As typical in (simultaneous) game theory, the optimal strategy is usually mixed, i.e., probabilistic, for both the attacker and the defender. From the point of view of information flow, this was to be expected in the case of the defender, since it is well known that randomization at the level of the system design may help to reduce information leaks. Regarding the attacker, however, this seems the first work (w.r.t. the literature in information flow) proving formally that in certain cases the optimal attack strategy is necessarily probabilistic.

6.1.2. Efficient Utility Improvement for Location Privacy

The continuously increasing use of location-based services poses an important threat to the privacy of users. A natural defense is to employ an obfuscation mechanism, such as those providing geo-indistinguishability [24], a framework for obtaining formal privacy guarantees that has become popular in recent years. Ideally, one would like to employ an optimal obfuscation mechanism, providing the best utility among those satisfying the required privacy level. In theory optimal mechanisms can be constructed via linear programming. In practice, however, this is only feasible for a radically small number of locations. As a consequence, all known applications of geo-indistinguishability simply use noise drawn from a planar Laplace distribution.

In [12], we studied methods for substantially improving the utility of location obfuscation, while maintaining practical applicability as a main goal. We provided such solutions for both infinite (continuous or discrete) as well as large but finite domains of locations, using a Bayesian remapping procedure as a key ingredient. We evaluated our techniques in two real world complete datasets, without any restriction on the evaluation area, and showed important utility improvements with respect to the standard planar Laplace approach.

6.1.3. Trading Optimality for Performance in Location Privacy

Location-Based Services (LBSs) provide invaluable aid in the everyday activities of many individuals, however they also pose serious threats to the user' privacy. There is, therefore, a growing interest in the development of mechanisms to protect location privacy during the use of LBSs. Nowadays, the most popular methods are probabilistic, and the so-called optimal method achieves an optimal trade-off between privacy and utility by using linear optimization techniques.

Unfortunately, due to the complexity of linear programming, the method is unfeasible for a large number N of locations, because the constraints are $O(N^3)$. In [20], we have proposed a technique to reduce the number of constraints to $O(N^2)$, at the price of renouncing to perfect optimality. We have showed however that on practical situations the utility loss is quite acceptable, while the gain in performance is significant.

6.1.4. Methods for Location Privacy: A comparative overview

The growing popularity of location-based services, allowing to collect huge amounts of information regarding users' location, has started raising serious privacy concerns. In [13] we analyzed the various kinds of privacy breaches that may arise in connection with the use of location-based services, and we surveyd and compared the metrics and the mechanisms that have been proposed in the literature.

6.1.5. Quantifying Leakage in the Presence of Unreliable Sources of Information

Belief and min-entropy leakage are two well-known approaches to quantify information flow in security systems. Both concepts stand as alternatives to the traditional approaches founded on Shannon entropy and mutual information, which were shown to provide inadequate security guarantees. In [16] we unified the two concepts in one model so as to cope with the frequent (potentially inaccurate, misleading or outdated) attackers' side information about individuals on social networks, online forums, blogs and other forms of online communication and information sharing. To this end we proposed a new metric based on min-entropy that takes into account the adversary's beliefs.

6.1.6. Differential Inference Testing: A Practical Approach to Evaluate Anonymized Data

In order to protect individuals' privacy, governments and institutions impose some obligations on data sharing and publishing. Mainly, they require the data to be "anonymized". In this paper, we have shortly discussed the criteria introduced by European General Data Protection Regulation to assess anonymized data. We have argued that the evaluation of anonymized data should be based on whether the data allows individual based inferences, instead of being centered around the concept of re-identification as the regulation has proposed.

Then, we have proposed an inference-based framework that can be used to evaluate the robustness of a given anonymized dataset against a specific inference model, e.g. a machine learning model.

Our approach evaluates the anonymized data itself, and deals with the related anonymization technique as a black-box. Thus, it can be used to assess datasets that are anonymized by organizations which may prefer not to provide access to their techniques. Finally, we have used our framework to evaluate two datasets after being anonymized using k-anonymity and l-diversity.

6.1.7. Formal Analysis and Offline Monitoring of Electronic Exams

More and more universities are moving toward electronic exams (in short e-exams). This migration exposes exams to additional threats, which may come from the use of the information and communication technology. In [17], we have identified and defined several security properties for e-exam systems. Then, we have showed how to use these properties in two complementary approaches: model-checking and monitoring.

We have illustrated the validity of our definitions by analyzing a real e-exam used at the pharmacy faculty of University Grenoble Alpes (UGA) to assess students. On the one hand, we have instantiated our properties as queries for ProVerif, a process calculus based automatic verifier for cryptographic protocols,

and we have used it to check our modeling of UGA exam specifications. ProVerif found some attacks. On the other hand, we have expressed our properties as Quantified Event Automata (QEAs), and we have synthesized them into monitors using MarQ, a Java tool designed to implement QEAs. Then, we have used these monitors to verify real exam executions conducted by UGA. Our monitors found fraudulent students and discrepancies between the specifications of UGA exam and its implementation.

6.1.8. On the Compositionality of Quantitative Information Flow

In the min-entropy approach to quantitative information flow, the leakage is defined in terms of a minimization problem, which, in the case of large systems, can be computationally rather heavy. The same happens for the recently proposed generalization called g-vulnerability. In [18] we studied the case in which the channel associated to the system can be decomposed into simpler channels, which typically happens when the observables consist of several components. Our main contribution is the derivation of bounds on the g-leakage of the whole system in terms of the g-leakages of its components. We also considered the particular cases of min-entropy leakage and of parallel channels, generalizing and systematizing results from the literature. We demonstrated the effectiveness of our method and evaluate the precision of our bounds using examples.

6.2. Foundations of Concurrency

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

6.2.1. Declarative Framework for Semantical Interpretations of Structured Information — An Applicative Approach.

Spatial constraint systems are algebraic structures from concurrent constraint programming to specify spatial and epistemic behavior in multi-agent system. In [21], [15] we studied the applicability of declarative models to encode and describe structured information by means of semantics. Specifically, we introduced D-SPACES, an implementation of constraint systems with space and extrusion operators. D-SPACES provides property-checking methods as well as an implementation of a specific type of constraint systems (a spatial boolean algebra). We showed the applicability of this framework with two examples; a scenario in the form of a social network where users post their beliefs and utter their opinions, and a semantical interpretation of a logical language to express time behaviors and properties.

6.2.2. Characterizing Right Inverses for Spatial Constraint Systems with Applications to Modal Logic

In [23] spatial constraint systems were used to give an abstract characterization of the notion of normality in modal logic and to derive right inverse/reverse operators for modal languages. In particular, a necessary and sufficient condition for the existence of right inverses was identified and the abstract notion of normality is shown to correspond to the preservation of finite suprema. Furthermore, a taxonomy of normal right inverses was provided, identifying the greatest normal right inverse as well as the complete family of minimal right inverses. These results were applied to existing modal languages such as the weakest normal modal logic, Hennessy-Milner logic, and linear-time temporal logic. Some implications of these results were also discussed in the context of modal concepts such as bisimilarity and inconsistency invariance.

6.2.3. Observational and Behavioural Equivalences for Soft Concurrent Constraint Programming

In citegadducci:hal-01675060 we presented a labelled semantics for Soft Concurrent Constraint Programming (SCCP), a meta-language where concurrent agents may synchronise on a shared store by either posting or checking the satisfaction of (soft) constraints. SCCP generalises the classical formalism by parametrising the constraint system over an order-enriched monoid, thus abstractly representing the store with an element of the monoid, and the standard unlabelled semantics just observes store updates. The novel operational rules were shown to offer a sound and complete co-inductive technique to prove the original equivalence over the unlabelled semantics. Based on this characterisation, we provided an axiomatisation for finite agents.

6.2.4. *On the Expressiveness of Spatial Constraint Systems*

The dissertation [11] focused on the expressiveness of spatial constraint systems in the broader perspective of modal and epistemic behaviour. It was shown that that spatial constraint systems are sufficiently robust to capture inverse modalities and to derive new results for modal logics. It was shown that one can use scs's to express a fundamental epistemic behaviour such as knowledge. The dissertation also provided an algebraic characterization of the notion of distributed information by means of constructors over scs's.

7. Partnerships and Cooperations

7.1. Regional Initiatives

7.1.1. *OPTIMEC*

Project title: Optimal Mechanisms for Privacy Protection

Funded by: DigiCosme

Duration: September 2016 - August 2019

Coordinator: Catuscia Palamidessi, Inria Saclay, EPI Comète

Other PI's: Serge Haddadm ENS Cachan.

Abstract: In this project we plan to investigate classes of utility and privacy measures, and to devise methods to obtain optimal mechanisms with respect to the trade-off between utility and privacy. In order to represent the probabilistic knowledge of the adversary and of the user, and the fact that mechanisms themselves can be randomized, we will consider a probabilistic setting. We will focus, in particular, on measures that are expressible as linear functions of the probabilities.

7.2. National Initiatives

7.2.1. *REPAS*

Program: ANR Blanc

Project title: Reliable and Privacy-Aware Software Systems via Bisimulation Metrics

Duration: October 2016 - September 2021

Coordinator: Catuscia Palamidessi, Inria Saclay, EPI Comète

Other PI's and partner institutions: Ugo del Lago, Inria Sophia Antipolis (EPI Focus) and University of Bologna (Italy). Vincent Danos, ENS Paris. Filippo Bonchi, ENS Lyon.

Abstract: In this project, we aim at investigating quantitative notions and tools for proving program correctness and protecting privacy. In particular, we will focus on bisimulation metrics, which are the natural extension of bisimulation on quantitative systems. As a key application, we will develop a mechanism to protect the privacy of users when their location traces are collected.

7.3. International Initiatives

7.3.1. *Inria Associate Teams*

7.3.1.1. *LOGIS*

Title: Logical and Formal Methods for Information Security

Inria principal investigator: Konstantinos Chatzikokolakis

International Partners:

Mitsuhiro Okada, Keio University (Japan)

Yusuke Kawamoto, AIST (Japan)

Tachio Terauchi, JAIST (Japan)

Masami Hagiya, University of Tokyo (Japan)

Start year: 2016

URL: http://www.lix.polytechnique.fr/~kostas/projects/logis/

Abstract: The project aims at integrating the logical / formal approaches to verify security protocols with (A) complexity theory and (B) information theory. The first direction aims at establishing the foundations of logical verification for security in the computational sense, with the ultimate goal of automatically finding attacks that probabilistic polynomial-time adversaries can carry out on protocols. The second direction aims at developing frameworks and techniques for evaluating and reducing information leakage caused by adaptive attackers.

7.3.2. Inria International Partners

7.3.2.1. Informal International Partners

Giovanni Cherubin, Royal Holloway, University of London, UK

Geoffrey Smith, Florida International University, USA

Carroll Morgan, NICTA , Australia

Annabelle McIver, Maquarie University, Australia

Moreno Falaschi, Professor, University of Siena, Italy

Mario Ferreira Alvim Junior, Assistant Professor, Federal University of Minas Gerais, Brazil

Camilo Rueda, Professor, Universidad Javeriana de Cali, Colombia

Carlos Olarte, Universidade Federal do Rio Grande do Norte, Brazil

Camilo Rocha, Associate Professor, Universidad Javeriana de Cali, Colombia

7.3.3. Participation in Other International Programs

7.3.3.1. CLASSIC

Program: Colciencias - Conv. 712.

Project acronym: CLASSIC.

Project title: Concurrency, Logic and Algebra for Social and Spatial Interactive Computation.

Duration: Oct 2016 - Oct 2019.

URL: http://goo.gl/Gv6Lij

Coordinator: Camilo Rueda, Universidad Javeriana de Cali, Colombia.

Other PI's and partner institutions: Carlos Olarte, Universidade Federal do Rio Grande do Norte, Brazil. Frank Valencia, CNRS-LIX and Inria Saclay.

Abstract: This project will advance the state of the art of domains such as mathematical logic, order theory and concurrency for reasoning about spatial and epistemic behaviour in multi-agent systems..

7.3.3.2. EPIC

Program: STIC-Amsud.

Project acronym: EPIC.

Project title: EPistemic Interactive Concurrency/

Duration: Oct 2016 - Oct 2019.

URL: https://sites.google.com/site/sticamsudepic/

Coordinator: Frank Valencia, CNRS-LIX and Inria Saclay.

Other PI's and partner institutions: Carlos Olarte, Universidade Federal do Rio Grande do Norte, Brazil. Camilo Rueda, Universidad Javeriana de Cali, Colombia.

Abstract: The aim of the project is to coherently combine and advance the state of the art of domains such as concurrency theory, information theory and rewriting systems for reasoning about social networks.

7.4. International Research Visitors

7.4.1. Visits of International Scientists

David de Frutos Escrig. Professor, Universidad Complutense Madrid, Spain. Jan-Feb 2017

Giovanni Cherubin, PhD student, Royal Holloway, University of London, UK. May 2017 and Oct 2017

Yusuke Kawamoto, Assistant Professor, National Institute of Advanced Industrial Science and Technology (AIST), Japan. July 2017 and Nov 2017

Carlos Olarte, Assistant Professor, Universidade Federal do Rio Grande do Norte, Brazil. July 2017

Camilo Rocha, Associate Professor, Universidad Javeriana de Cali, Colombia. Oct 2017

Camilo Rueda, Professor, Universidad Javeriana de Cali, Colombia. Nov 2017

Mario Ferreira Alvim Junior, Assistant Professor, Federal University of Minas Gerais, Brazil. Dec 2017

7.4.2. Internships

Anna Pazii. Univ. of Kiev, Ukraine. From July 2016 until Jan 2017.

Hector Delgado, Universidad Javeriana de Cali, Colombia. From May 2017 until July 2017.

Marco Romanelli. Univ. of Siena, Italy. From June 2017 until Sept 2017.

Georgi Dikov. Tech. Univ. of Munich, Germany. From Sept 2017 until Nov 2017.

Joaquin Felici. Univ. of Cordoba, Argentina. From Sept 2017 until Jan 2018.

Santiago Quintero, Universidad Javeriana de Cali, Colombia. From Nov until Dec 2017.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific events organisation

8.1.1.1. Member of the organizing committee

Catuscia Palamidessi is member of:

The Executive Committee of SIGLOG, the ACM Special Interest Group on Logic and Computation. Since 2014.

The Organizing Committee of LICS, the ACM/IEEE Symposium on Logic in Computer Science. Since 2010.

The Steering Committee of ETAPS, the European Joint Conferences on Theory and Practice of Software. Since 2006.

The Steering Committee of EACSL, the European Association for Computer Science Logics. Since 2015.

The Steering Committee of CONCUR, the International Conference in Concurrency Theory. Since 2016.

The Steering Committee of FORTE, the International Conference on Formal Techniques for Distributed Objects, Components, and Systems. Since 2014.

The IFIP Technical Committee 1 – Foundations of Computer Science. Since 2007.

The IFIP Working Group 1.7 – Theoretical Foundations of Security Analysis and Design. Since 2010.

The IFIP Working Group 1.8 – Concurrency Theory.

Frank D. Valencia is member of:

The steering committee of the International Workshop in Concurrency EXPRESS. Since 2010.

8.1.2. Scientific events selection

8.1.2.1. Member of conference program committees

Catuscia Palamidessi is/has been a member of the program committees of the following conferences and workshops:

PETS 2019. The 19th Privacy Enhancing Technologies Symposium. July 2019.

TASE 2018. The 12th International Symposium on Theoretical Aspects of Software Engineering Guangzhou, China, 29-31 August 2018.

PETS 2018. The 18th Privacy Enhancing Technologies Symposium. Barcelona, Spain, 24-27 July 2018.

FOSSACS 2018. The 21st International Conference on Foundations of Software Science and Computation Structures. (Part of ETAPS 2018.) Thessaloniki, Greece, 14-21 April 2018.

SOFSEM 2018. The 44th Annual Int'l Conference on Current Trends in Theory and Practice of Computer Science (track on Foundations of Computer Science). Krems an der Donau, Austria, 29 January- 2 February, 2018.

ICTAC 2017. The 14th International Colloquium on Theoretical Aspects of Computing. Hanoi, Vietnam, 23-27 October 2017.

TASE 2017. The 11th International Symposium on Theoretical Aspects of Software Engineering. Nice, France, 13-15 September 2017.

CONCUR 2017. The 28th International Conference on Concurrency Theory. Berlin, Germany, 5-8 September 2017.

CSL 2017. The 26th EACSL Annual Conference on Computer Science Logic. Stockholm, Sweden, 20-25 August 2017.

ICSOFT-PT 2017. The 12th International Conference on Software Paradigm Trends. Lisbon, Portugal, 24-26 July 2017.

ICALP 2017 (Track B). The 44th International Colloquium on Automata, Languages, and Programming. Warsaw, Poland, 10–14 July 2017.

FORTE 2017. The 37th IFIP International Conference on Formal Techniques for Distributed Objects, Components, and Systems. Neuchâtel, Switzerland, 19-22 June 2017.

CSR 2017. The 12th International Computer Science Symposium in Russia. Kazan, Russia, 8–12 June 2017.

Konstantinos Chatzikokolakis is/has been a member of the program committees of the following conferences and workshops:

BMDA 2018: Workshop on Big Mobility Data Analytics

QAPL 2018: International Workshop on Quantitative Aspects of Programming Languages and Systems

HotSpot 2018: 6th Workshop on Hot Issues in Security Principles and Trust

ICDE 2017: IEEE International Conference on Data Engineering

CSF 2017: 30th IEEE Computer Security Foundations Symposium

POST 2017: 6th International Conference on Principles of Security and Trust

BIGQP 2017: International Workshop on Big Geo Data Quality and Privacy

Frank D. Valencia is/has been a member of the program committees of the following conferences and workshops:

RADICAL-2017. International Workshop Recent Advances in Concurrency and Logic - RADICAL

CP-ICLP-SAT-DP-17. Doctoral Program of the 23rd International Conference on Principles and Practice of Constraint Programming

8.1.2.2. Reviewing

The members of the team reviewed several papers for international conferences and workshops.

8.1.3. Journals

8.1.3.1. Member of the editorial board

Catuscia Palamidessi is:

Member of the Editorial Board of Proceedings on Privacy Enhancing Technologies (PoPETs), published by De Gruyter.

Member of the Editorial Board of Mathematical Structures in Computer Science, published by the Cambridge University Press.

Member of the Editorial Board of Acta Informatica, published by Springer.

Member of the Editorial Board of the Electronic Notes of Theoretical Computer Science, published by Elsevier Science.

Member of the Editorial Board of LIPIcs: Leibniz International Proceedings in Informatics, Schloss Dagstuhl–Leibniz Center for Informatics.

Konstantinos Chatzikokolakis is:

Editorial board member of the newly established Proceedings on Privacy Enhancing Technologies (PoPETs), a scholarly journal for timely research papers on privacy.

8.1.3.2. Reviewing

The members of the team regularly review papers for international journals and conferences.

8.1.4. Other Editorial Activities

Frank D. Valencia has been:

Co-editor of the special issue on Mathematical Structures in Computer Science dedicated to the best papers from the 12th International Colloquium on Theoretical Aspects of Computing.

8.1.5. Participation in other committees

Catuscia Palamidessi has been serving in the following committees:

Member of the committee for the assignment of the Inria International Chairs.

Member of the committee for the Alonzo Church Award for Outstanding Contributions to Logic and Computation. Since 2015. In 2018 Palamidessi is the president of this committee.

President of the selection committee for the EATCS Best Paper Award at the ETAPS conferences. Since 2006.

8.1.6. Invited talks

Catuscia Palamidessi has given invited talks at the following conferences and workshops:

FACS 2017. The 14th International Conference on Formal Aspects of Component Software. Barga, Portugal. 10-13 Oct, 2017.

Cybersecurity 2017. Focus Day on Cyber Security and Helthcare. In the context of the European Cyber Week. Rennes, France, 30 November 2017.

QuaSy 2017. Quantitative Systems: Theory and Applications. Como, Italy, 16-17 October 2017.

Women in Logic 2017, Reykjavik, Island, June 2017.

CrossFyre 2017 Workshop on Cryptography, Robustness, and Provably Secure Schemes. Paris. April 2017.

FORSE 2017 (Keynote speaker). 1st International Workshop on FORmal methods for Security Engineering. Porto, Portugal. 19–21 February, 2017.

8.1.7. Service

Catuscia Palamidessi has served as:

Reviewer for the projects proposal for the program PRIN, sponsored by the Italian MIUR ("Ministero dell'Istruzione, dell'Università e della Ricerca"). Since 2004.

Frank Valencia has served as:

Directeur adjoint de l'UMR 7161, le Laboratoire d'Informatique de l'Ecole Polytechnique (LIX). May 2016 - .

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Master : Frank D. Valencia has been teaching the undergraduate course "Computability", 45 hours, at the Pontificia Universidad Javeriana de Cali, Colombia. July 27 - Nov 1, 2017.

Master : Frank D. Valencia has been teaching the masters course "Foundations of Computer Science", 45 hours, at the Pontificia Universidad Javeriana de Cali, Colombia. Jan 27 - Jun 1, 2017.

Master: Konstantinos Chatzikokolakis and Catuscia Palamidessi have been teaching a course on the Foundations of Privacy at the MPRI, the Master Parisien pour la Recherche en Informatique. University of Paris VII. A.Y. 2016-17 and 2017-18. Total for each semester: 24 hours plus 6 hours for the exam and the exercise session is preparation to the exam.

8.2.2. Supervision

PhD in progress (2017-) Marco Romanelli. Co-supervised by Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Moreno Falaschi (University of Siena, Italy). Thesis subject: Application of Information Flow to feature selection in machine learning.

PhD in progress (2017-) Anna Pazii. Co-supervised by Konstantinos Chatzikokolakis and Catuscia Palamidessi. Thesis subject: Local Differential Privacy.

PhD in progress (2016-) Tymofii Prokopenko. Ecole Polytechnique and ENS Cachan. Grant Digiteo-Digicosme. Co-supervised by Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Serge Haddad (ENS Cachan).

PhD in progress (2017-) Sergio Ramirez. Co-supervised by Frank Valencia and Camilo Rueda, Universidad Javeriana Cali. Thesis subject: Quantitive Spatial Constraint Systems.

PhD terminated (2015-17) Joris Lamare. Ecole Polytechnique. Grant MSR Center. Co-supervised by Catuscia Palamidessi and Konstantinos Chatzikokolakis. Joris has stopped his PhD due to personal reasons.

PhD completed (2014-17) Michel Guzman. Titile: On the Expressiveness of Spatial Constraint Systems [11]. Ecole Polytechnique. Grant Inria CORDI-S. Co-supervised by Catuscia Palamidessi and Frank D. Valencia.

8.2.3. Juries

Catuscia Palamidessi has been reviewer and member of the board at the PhD defense for the thesis of the following PhD student:

Nicolas Bonifas (Ecole Polytechnique, France). Member of the committee board at the PhD defense. Title of the thesis: *Geometric and Dual Approaches to Cumulative Scheduling*. Supervised by Philippe Baptiste. Defended in December 2017.

Maggie Mhanna (CentraleSupelec, France). Member of the committee board at the PhD defense. Supervised by Pablo Piantanida. Defended in January 2017.

8.2.4. Other didactical duties

Catuscia Palamidessi is:

Member of the advising committee for Hamid Ebadi, PhD student supervised by David Sands, Chalmers University, Sweden, since 2014. Also reviewer and member of the committee for the half-way thesis defense (Licentiate) that took place in June 2015.

External member of the scientific council for the PhD in Computer Science at the University of Pisa, Italy. Since 2012.

Member of the advising committee for the PhD of Jun Wang (PhD student supervised by Qiang Tang and Peter Ryan), University of Luxembourg. Since December 2014.

Member of the advising committee for the PhD of Andrea Margheri (PhD student supervised by Rosario Pugliese), University of Florence, Italy. 2014-16.

Konstantinos Chatzikokolakis and Catuscia Palamidessi have designed, and coordinate, a course on the Foundations of Privacy at the MPRI, the Master Parisien pour la Recherche en Informatique. University of Paris VII. A.Y. Since 2015.

9. Bibliography

Major publications by the team in recent years

[1] M. S. ALVIM, M. E. ANDRÉS, K. CHATZIKOKOLAKIS, P. DEGANO, C. PALAMIDESSI.*On the information leakage of differentially-private mechanisms*, in "Journal of Computer Security", 2015, vol. 23, n^o 4, p. 427-469 [*DOI :* 10.3233/JCS-150528], https://hal.inria.fr/hal-00940425.

[2] M. S. ALVIM, K. CHATZIKOKOLAKIS, A. MCIVER, C. MORGAN, C. PALAMIDESSI, G. SMITH.*Additive and multiplicative notions of leakage, and their capacities*, in "27th Computer Security Foundations Symposium (CSF 2014)", Vienna, Austria, IEEE, July 2014, p. 308–322 [*DOI :* 10.1109/CSF.2014.29], https://hal.inria.fr/hal-00989462.

[3] M. S. ALVIM, K. CHATZIKOKOLAKIS, A. MCIVER, C. MORGAN, C. PALAMIDESSI, G. SMITH.*Axioms for Information Leakage*, in "29th Computer Security Foundations Symposium (CSF 2016)", Lisbon, Portugal, IEEE, June 2016, 16, https://hal.inria.fr/hal-01330414.

[4] M. ALVIM, K. CHATZIKOKOLAKIS, C. PALAMIDESSI, G. SMITH.*Measuring Information Leakage using Generalized Gain Functions*, in "Computer Security Foundations", Cambridge MA, United States, IEEE, 2012, p. 265-279 [*DOI :* 10.1109/CSF.2012.26], http://hal.inria.fr/hal-00734044.

[5] M. ANDRÉS, N. BORDENABE, K. CHATZIKOKOLAKIS, C. PALAMIDESSI.*Geo-Indistinguishability: Differential Privacy for Location-Based Systems*, in "20th ACM Conference on Computer and Communications Security", Berlin, Allemagne, ACM Press, 2013, p. 901-914, DGA, Inria large scale initiative CAPPRIS [*DOI :* 10.1145/2508859.2516735], http://hal.inria.fr/hal-00766821.

[6] A. ARISTIZÁBAL, F. BONCHI, C. PALAMIDESSI, L. PINO, D. VALENCIA.*Deriving Labels and Bisimilarity for Concurrent Constraint Programming*, in "FOSSACS 2011 : 14th International Conference on Foundations of Software Science and Computational Structures", Saarbrücken, Germany, M. HOFMANN (editor), Lecture Notes in Computer Science, Springer, March 2011, vol. 6604, p. 138-152 [*DOI :* 10.1007/ISBN 978-3-642-19804-5], https://hal.archives-ouvertes.fr/hal-00546722.

[7] N. E. BORDENABE, K. CHATZIKOKOLAKIS, C. PALAMIDESSI.*Optimal Geo-Indistinguishable Mechanisms for Location Privacy*, in "CCS - 21st ACM Conference on Computer and Communications Security", Scottsdale, Arizona, United States, G.-J. AHN, M. YUNG, N. LI (editors), Proceedings of the 21st ACM Conference on Computer and Communications Security, ACM, November 2014, p. 251-262 [*DOI :* 10.1145/2660267.2660345], https://hal.inria.fr/hal-00950479.

[8] K. CHATZIKOKOLAKIS, C. PALAMIDESSI, M. STRONATI.*Constructing elastic distinguishability metrics for location privacy*, in "Proceedings on Privacy Enhancing Technologies", June 2015, vol. 2015, n^{o} 2, p. 156-170 [*DOI :* 10.1515/POPETS-2015-0023], https://hal.inria.fr/hal-01270197.

[9] M. GUZMÁN, S. HAAR, S. PERCHY, C. RUEDA, F. VALENCIA.*Belief, Knowledge, Lies and Other Utterances in an Algebra for Space and Extrusion*, in "Journal of Logical and Algebraic Methods in Programming", September 2016 [*DOI :* 10.1016/J.JLAMP.2016.09.001], https://hal.inria.fr/hal-01257113.

[10] S. KNIGHT, C. PALAMIDESSI, P. PANANGADEN, F. D. VALENCIA.*Spatial and Epistemic Modalities in Constraint-Based Process Calculi*, in "CONCUR 2012 - Concurrency Theory - 23rd International Conference, CONCUR 2012", Newcastle upon Tyne, United Kingdom, September 2012, vol. 7454, p. 317-332 [*DOI :* 10.1007/978-3-642-32940-1], http://hal.inria.fr/hal-00761116.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] M. GUZMÁN.*On the Expressiveness of Spatial Constraint Systems*, École Polytechnique X, September 2017, https://hal.inria.fr/tel-01674956.

Articles in International Peer-Reviewed Journal

[12] K. CHATZIKOKOLAKIS, E. ELSALAMOUNY, C. PALAMIDESSI.*Efficient Utility Improvement for Location Privacy*, in "Proceedings on Privacy Enhancing Technologies", 2017, vol. 2017, n^{o} 4, p. 308-328 [*DOI :* 10.1515/POPETS-2017-0051], https://hal.inria.fr/hal-01422842.

[13] K. CHATZIKOKOLAKIS, E. ELSALAMOUNY, C. PALAMIDESSI, A. PAZII.*Methods for Location Privacy: A comparative overview*, in "Foundations and Trends® in Privacy and Security ", 2017, vol. 1, n^{o} 4, p. 199-257, Submitted for publication, https://hal.inria.fr/hal-01421457.

[14] F. GADDUCCI, F. SANTINI, L. F. PINO DUQUE, F. VALENCIA.*Observational and Behavioural Equivalences for Soft Concurrent Constraint Programming*, in "Journal of Logical and Algebraic Methods in Programming", November 2018, vol. 92, p. 45-63 [*DOI :* 10.1016/J.JLAMP.2017.06.001], https://hal.inria.fr/hal-01675060.

[15] S. HAAR, S. PERCHY, F. VALENCIA.*Declarative Framework for Semantical Interpretations of Structured Information — An Applicative Approach*, in "International Journal of Semantic Computing", December 2017, vol. 11, n^{o} 04, p. 451 - 472 [*DOI :* 10.1142/S1793351X17400189], https://hal.inria.fr/hal-01673529.

[16] S. HAMADOU, C. PALAMIDESSI, V. SASSONE.*Quantifying Leakage in the Presence of Unreliable Sources of Information*, in "Journal of Computer and System Sciences", 2017, vol. 88, p. 27-52, To appear, https://hal.inria.fr/hal-01421417.

[17] A. KASSEM, Y. FALCONE, P. LAFOURCADE.*Formal analysis and offline monitoring of electronic exams*, in "Formal Methods in System Design", August 2017, vol. 51, n^{o} 1, p. 117 - 153 [*DOI :* 10.1007/s10703-017-0280-0], https://hal.inria.fr/hal-01653884.

[18] Y. KAWAMOTO, K. CHATZIKOKOLAKIS, C. PALAMIDESSI.*On the Compositionality of Quantitative Information Flow*, in "Logical Methods in Computer Science", August 2017, vol. 13, n^{o} 3, p. 1-31, https://arxiv.org/abs/1611.00455 - Submitted for publication to Logical Methods in Computer Science, https://hal.inria.fr/hal-01421424.

International Conferences with Proceedings

[19] M. S. ALVIM, K. CHATZIKOKOLAKIS, Y. KAWAMOTO, C. PALAMIDESSI. *Information Leakage Games*, in "Decision and Game Theory for Security - 8th International Conference", Vienna, Austria, S. RASS, B. AN, C. KIEKINTVELD, F. FANG, S. SCHAUER (editors), Lecture Notes in Computer Science, Springer, October 2017, vol. 10575, p. 437-457 [*DOI :* 10.1007/978-3-319-68711-7_23], https://hal.inria.fr/hal-01678950.

[20] K. CHATZIKOKOLAKIS, S. HADDAD, A. KASSEM, C. PALAMIDESSI. *Trading Optimality for Performance in Location Privacy*, in "11th EAI International Conference on Performance Evaluation Methodologies and Tools", Venice, Italy, December 2017, https://hal.inria.fr/hal-01678256.

[21] S. HAAR, S. PERCHY, F. VALENCIA.*D-SPACES: Implementing Declarative Semantics for Spatially Structured Information*, in "11th International Conference on Semantic Computing", San Diego, California, United States, IEEE ICSC 2017, IEEE, January 2017, vol. 11, https://hal.inria.fr/hal-01328189.

Research Reports

[22] A. KASSEM, G. ACS, C. CASTELLUCCIA.*Differential Inference Testing A Practical Approach to Evaluate Anonymized Data*, Inria, January 2018, https://hal.inria.fr/hal-01681014.

Other Publications

[23] M. GUZMÁN, S. PERCHY, C. RUEDA, F. VALENCIA.*Characterizing Right Inverses for Spatial Constraint Systems with Applications to Modal Logic* , January 2018, Submitted to Theoretical Computer Science (TCS), https://hal.inria.fr/hal-01675010.

References in notes

[24] M. ANDRÉS, N. BORDENABE, K. CHATZIKOKOLAKIS, C. PALAMIDESSI.*Geo-Indistinguishability: Differential Privacy for Location-Based Systems*, in "20th ACM Conference on Computer and Communications Security", Berlin, Allemagne, ACM Press, 2013, p. 901-914, DGA, Inria large scale initiative CAPPRIS [*DOI :* 10.1145/2508859.2516735], http://hal.inria.fr/hal-00766821.

Project-Team COMMANDS

Control, Optimization, Models, Methods and Applications for Nonlinear Dynamical Systems

IN COLLABORATION WITH: Centre de Mathématiques Appliquées (CMAP), Unité de Mathématiques Appliquées (UMA - ENSTA)

IN PARTNERSHIP WITH:
CNRS

Ecole Polytechnique

RESEARCH CENTER
Saclay - Île-de-France

THEME
Optimization and control of dynamic systems

Table of contents

Project-Team COMMANDS

Creation of the Project-Team: 2009 January 01

Keywords:

Computer Science and Digital Science:

A6.2.1. - Numerical analysis of PDE and ODE
A6.2.6. - Optimization
A6.2.7. - High performance computing
A6.3.2. - Data assimilation
A6.4.1. - Deterministic control
A6.4.2. - Stochastic control

Other Research Topics and Application Domains:

B4.4.1. - Smart grids
B7.1.2. - Road traffic
B7.1.3. - Air traffic
B7.2.1. - Smart vehicles

1. Personnel

Research Scientists

Joseph Frederic Bonnans [Team leader, Inria, Senior Researcher, HDR]
Axel Kroner [Inria, Starting Research Position, until Mar 2017]
Pierre Martinon [Inria, Researcher]

External Collaborator

Axel Kroner [U. Humboldt, from Apr 2017]

Technical Staff

Jinyan Liu [Inria]

PhD Students

Arthur Le Rhun [Ifpen]
Cedric Rommel [Safety Line]

Visiting Scientist

Justina Gianatti [U. Rosario, from Nov 2017]

Administrative Assistants

Hanadi Dib [Inria]
Jessica Gameiro [Inria]

2. Overall Objectives

2.1. Scientific directions

Commands is a team devoted to dynamic optimization, both for deterministic and stochastic systems. This includes the following approaches: trajectory optimization, deterministic and stochastic optimal control, stochastic programming, dynamic programming and Hamilton-Jacobi-Bellman equation.

Our aim is to derive new and powerful algorithms for solving numerically these problems, with applications in several industrial fields. While the numerical aspects are the core of our approach it happens that the study of convergence of these algorithms and the verification of their well-posedness and accuracy raises interesting and difficult theoretical questions, such as, for trajectory optimization: qualification conditions and second-order optimality condition, well-posedness of the shooting algorithm, estimates for discretization errors; for the Hamilton-Jacobi-Bellman approach: accuracy estimates, strong uniqueness principles when state constraints are present, for stochastic programming problems: sensitivity analysis.

2.2. Industrial impact

For many years the team members have been deeply involved in various industrial applications, often in the framework of PhD theses. The Commands team itself has dealt since its foundation in 2009 with several types of applications:

- Space vehicle trajectories, in collaboration with CNES, the French space agency.
- Aeronautics, in collaboration with the startup Safety Line.
- Production, management, storage and trading of energy resources, in collaboration with EDF, GDF and TOTAL.
- Energy management for hybrid vehicles, in collaboration with Renault and IFPEN.

We give more details in the Bilateral contracts section.

3. Research Program

3.1. Historical aspects

The roots of deterministic optimal control are the "classical" theory of the calculus of variations, illustrated by the work of Newton, Bernoulli, Euler, and Lagrange (whose famous multipliers were introduced in [32]), with improvements due to the "Chicago school", Bliss [24] during the first part of the 20th century, and by the notion of relaxed problem and generalized solution (Young [37]).

Trajectory optimization really started with the spectacular achievement done by Pontryagin's group [36] during the fifties, by stating, for general optimal control problems, nonlocal optimality conditions generalizing those of Weierstrass. This motivated the application to many industrial problems (see the classical books by Bryson and Ho [28], Leitmann [34], Lee and Markus [33], Ioffe and Tihomirov [31]).

Dynamic programming was introduced and systematically studied by R. Bellman during the fifties. The HJB equation, whose solution is the value function of the (parameterized) optimal control problem, is a variant of the classical Hamilton-Jacobi equation of mechanics for the case of dynamics parameterized by a control variable. It may be viewed as a differential form of the dynamic programming principle. This nonlinear first-order PDE appears to be well-posed in the framework of *viscosity solutions* introduced by Crandall and Lions [29]. The theoretical contributions in this direction did not cease growing, see the books by Barles [22] and Bardi and Capuzzo-Dolcetta [21].

3.2. Trajectory optimization

The so-called *direct methods* consist in an optimization of the trajectory, after having discretized time, by a nonlinear programming solver that possibly takes into account the dynamic structure. So the two main problems are the choice of the discretization and the nonlinear programming algorithm. A third problem is the possibility of refinement of the discretization once after solving on a coarser grid.

In the *full discretization approach*, general Runge-Kutta schemes with different values of control for each inner step are used. This allows to obtain and control high orders of precision, see Hager [30], Bonnans [25]. In the *indirect* approach, the control is eliminated thanks to Pontryagin's maximum principle. One has then to solve the two-points boundary value problem (with differential variables state and costate) by a single or multiple shooting method. The questions are here the choice of a discretization scheme for the integration of the boundary value problem, of a (possibly globalized) Newton type algorithm for solving the resulting finite dimensional problem in IR^n (n is the number of state variables), and a methodology for finding an initial point.

3.3. Hamilton-Jacobi-Bellman approach

This approach consists in calculating the value function associated with the optimal control problem, and then synthesizing the feedback control and the optimal trajectory using Pontryagin's principle. The method has the great particular advantage of reaching directly the global optimum, which can be very interesting when the problem is not convex.

Optimal stochastic control problems occur when the dynamical system is uncertain. A decision typically has to be taken at each time, while realizations of future events are unknown (but some information is given on their distribution of probabilities). In particular, problems of economic nature deal with large uncertainties (on prices, production and demand). Specific examples are the portfolio selection problems in a market with risky and non-risky assets, super-replication with uncertain volatility, management of power resources (dams, gas). Air traffic control is another example of such problems.

For solving stochastic control problems, we studied the so-called Generalized Finite Differences (GFD), that allow to choose at any node, the stencil approximating the diffusion matrix up to a certain threshold [27]. Determining the stencil and the associated coefficients boils down to a quadratic program to be solved at each point of the grid, and for each control. This is definitely expensive, with the exception of special structures where the coefficients can be computed at low cost. For two dimensional systems, we designed a (very) fast algorithm for computing the coefficients of the GFD scheme, based on the Stern-Brocot tree [26].

4. Application Domains

4.1. Fuel saving by optimizing airplanes trajectories

We have a collaboration with the startup Safety Line on the optimization of trajectories for civil aircrafts. Key points include the reliable identification of the plane parameters (aerodynamic and thrust models) using data from the flight recorders, and the robust trajectory optimization of the climbing and cruise phases. We use both local (quasi-Newton interior-point algorithms) and global optimization tools (dynamic programming). The local method for the climb phase is in production and has been used for several hundreds of actual plane flights.

4.2. Hybrid vehicles

We have a collaboration with IFPEN on the energy management for hybrid vehicles. A significant direction is the analysis and classification of traffic data. More specifically, we focus on the traffic probability distribution in the (speed,torque) plane, with a time / space subdivision (road segments and timeframes).

4.3. Biological systems

We renewed in 2017 our interest in (micro)biological systems, joining projects Cosy and Algae in silico on the topic of the optimization of micro-organisms populations.

5. Highlights of the Year

5.1. Suboptimal feedback control of PDEs

In [13], J. Garcke (SCAI-Fraunhofer I.) and A. Kröner were able to solve finite time horizon suboptimal feedback control problems for partial differential equations is proposed by solving dynamic programming equations on adaptive sparse grids. The approach is illustrated for the wave equation and an extension to equations of Schrödinger type is discussed. A semi-discrete optimal control problem is introduced and the feedback control is derived from the corresponding value function. A semi-Lagrangian scheme is combined with spatially adaptive sparse grids. An adaptive grid refinement procedure is explored. We present several numerical examples studying the effect the parameters characterizing the sparse grid have on the accuracy of the value function and the optimal trajectory. Problems with dimensions up to eight were solved.

6. New Software and Platforms

6.1. BOCOP

Boite à Outils pour le Contrôle OPtimal
KEYWORDS: Dynamic Optimization - Identification - Biology - Numerical optimization - Energy management - Transportation
FUNCTIONAL DESCRIPTION: Bocop is an open-source toolbox for solving optimal control problems, with collaborations with industrial and academic partners. Optimal control (optimization of dynamical systems governed by differential equations) has numerous applications in transportation, energy, process optimization, energy and biology. Bocop includes a module for parameter identification and a graphical interface, and runs under Linux / Windows / Mac.

RELEASE FUNCTIONAL DESCRIPTION: Handling of delay systems Alternate automatic differentiation tool: CppAD Update for CMake and MinGW (windows version)

- Participants: Benjamin Heymann, Virgile Andreani, Jinyan Liu, Joseph Frédéric Bonnans and Pierre Martinon
- Contact: Pierre Martinon
- URL: http://bocop.org

6.2. Bocop HJB

KEYWORDS: Optimal control - Stochastic optimization - Global optimization
FUNCTIONAL DESCRIPTION: Toolbox for stochastic or deterministic optimal control, dynamic programming / HJB approach.

RELEASE FUNCTIONAL DESCRIPTION: User interface State jumps for switched systems Explicit handling of final conditions Computation of state probability density (fiste step to mean field games)

- Participants: Benjamin Heymann, Jinyan Liu, Joseph Frédéric Bonnans and Pierre Martinon
- Contact: Joseph Frédéric Bonnans
- URL: http://bocop.org

6.3. Bocop Avion

KEYWORDS: Optimization - Aeronautics
FUNCTIONAL DESCRIPTION: Optimize the climb speeds and associated fuel consumption for the flight planning of civil airplanes.

NEWS OF THE YEAR: Improved atmosphere model 2D interpolations for temperature and wind data

- Participants: Gregorutti Baptiste, Cindie Andrieu, Anamaria Lupu, Joseph Frédéric Bonnans, Karim Tekkal, Pierre Jouniaux and Pierre Martinon
- Partner: Safety Line
- Contact: Pierre Martinon
- URL: http://www.safety-line.fr

6.4. Bocop HJB Avion

KEYWORDS: Optimization - Aeronautics
FUNCTIONAL DESCRIPTION: Optimize the climb and cruising trajectory of flight by a HJB approach.
NEWS OF THE YEAR: First demonstrator for cruise flight deployed at Safety Line

- Participants: Pierre Martinon, Joseph Frédéric Bonnans, Jinyan Liu, Gregorutti Baptiste and Anamaria Lupu
- Partner: Safety Line
- Contact: Pierre Martinon
- URL: http://www.safety-line.fr

7. New Results

7.1. Deterministic Optimal Control

7.1.1. Galerkin approximations of nonlinear optimal control problems in Hilbert spaces

Participant: Axel Kroner.

With Mickaël D. Chekroun (UCLA), and Honghu Liu (Virginia Tech). Nonlinear optimal control problems in Hilbert spaces are considered for which we derive approximation theorems for Galerkin approximations. Approximation theorems are available in the literature. The originality of our approach relies on the identification of a set of natural assumptions that allows us to deal with a broad class of nonlinear evolution equations and cost functionals for which we derive convergence of the value functions associated with the optimal control problem of the Galerkin approximations. This convergence result holds for a broad class of nonlinear control strategies as well. In particular, we show that the framework applies to the optimal control of semilinear heat equations posed on a general compact manifold without boundary. The framework is then shown to apply to geoengineering and mitigation of greenhouse gas emissions formulated for the first time in terms of optimal control of energy balance climate models posed on the sphere S^2. See [12].

7.1.2. Galerkin approximations for the optimal control of nonlinear delay differential equations

Participant: Axel Kroner.

With Mickaël D. Chekroun (UCLA), and Honghu Liu (Virginia Tech).

Optimal control problems of nonlinear delay differential equations (DDEs) are considered for which we propose a general Galerkin approximation scheme built from Koornwinder polynomials. Error estimates for the resulting Galerkin-Koornwinder approximations to the optimal control and the value function, are derived for a broad class of cost function-als and nonlinear DDEs. The approach is illustrated on a delayed logistic equation set not far away from its Hopf bifurcation point in the parameter space. In this case, we show that low-dimensional controls for a standard quadratic cost functional can be efficiently computed from Galerkin-Koornwinder approximations to reduce at a nearly optimal cost the oscillation amplitude displayed by the DDE's solution. Optimal controls computed from the Pontryagin's maximum principle (PMP) and the Hamilton-Jacobi-Bellman equation (HJB) associated with the corresponding ODE systems, are shown to provide numerical solutions in good agreement. It is finally argued that the value function computed from the corresponding reduced HJB equation provides a good approximation of that obtained from the full HJB equation. See [16].

7.2. Stochastic Control

7.2.1. On the time discretization of stochastic optimal control problems: the dynamic programming approach

Participant: Frederic Bonnans.

With Justina Gianatti (U. Rosario) and Francisco J. Silva (U. Limoges) In this work we consider the time discretization of stochastic optimal control problems. Under general assumptions on the data, we prove the convergence of the value functions associated with the discrete time problems to the value function of the original problem. Moreover, we prove that any sequence of optimal solutions of discrete problems is minimizing for the continuous one. As a consequence of the Dynamic Programming Principle for the discrete problems, the minimizing sequence can be taken in discrete time feedback form. See [17].

7.2.2. Variational analysis for options with stochastic volatility and multiple factors

Participants: Frederic Bonnans, Axel Kroner.

We perform a variational analysis for a class of European or American options with stochastic volatility models, including those of Heston and Achdou-Tchou. Taking into account partial correlations and the presence of multiple factors, we obtain the well-posedness of the related partial differential equations, in some weighted Sobolev spaces. This involves a generalization of the commutator analysis introduced by Achdou and Tchou. See [18].

7.2.3. Infinite Horizon Stochastic Optimal Control Problems with Running Maximum Cost

Participant: Axel Kroner.

With Athena Picarelli (U. Oxford) and Hasna Zidani (ENSTA).

An infinite horizon stochastic optimal control problem with running maximum cost is considered. The value function is characterized as the viscosity solution of a second-order HJB equation with mixed boundary condition. A general numerical scheme is proposed and convergence is established under the assumptions of consistency, monotonicity and stability of the scheme. A convergent semi-Lagrangian scheme is presented in detail. See [19].

7.3. Applications

7.3.1. On the Design of Optimal Health Insurance Contracts under Ex Post Moral Hazard

Participant: Pierre Martinon.

With Pierre Picard and Anasuya Raj (Ecole Polytechnique, Econ. dpt).
We analyze in [20] the design of optimal medical insurance under ex post moral hazard, i.e., when illness severity cannot be observed by insurers and policyholders decide on their health expenditures. We characterize the trade-off between ex ante risk sharing and ex post incentive compatibility, in an optimal revelation mechanism under hidden information and risk aversion.

We establish that the optimal contract provides partial insurance at the margin, with a deductible when insurers rates are affected by a positive loading, and that it may also include an upper limit on coverage. We show that the potential to audit the health state leads to an upper limit on out-of-pocket expenses. Numerical simulations indicate that these qualitative results tend to be robust with respect to the health parameter.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. Safety Line

In the framework of an Ilab with startup Safety Line (http://www.safety-line.fr), we design tools for the optimization of fuel consumption for civil planes. A first part is devoted to the identification of the aerodynamic and thrust characteristics of the plane, using recorded data from hundreds of flights. As an illustration, Fig. 1 shows the drag and lift coefficients for a Boeing 737, as functions of Mach and angle of attack. Latest results have been presented by Cedric Rommel at [15].

A second part is optimizing the fuel consumption during the climb and cruise phases. Fig. 2 shows a simulated climb phase, along with recorded data from the actual flight. This collaboration relies significantly on the toolboxes BOCOP and BOCOPHJB developed by Commands since 2010. The resulting commercial tool OptiClimb is currently under testing in several airplane companies, totalling about a hundred actual optimized flights per day. Recent improvements include better atmosphere models and more accurate data for temperature and wind, as well as a first demonstrator for cruise flight optimization, see Fig. 3.

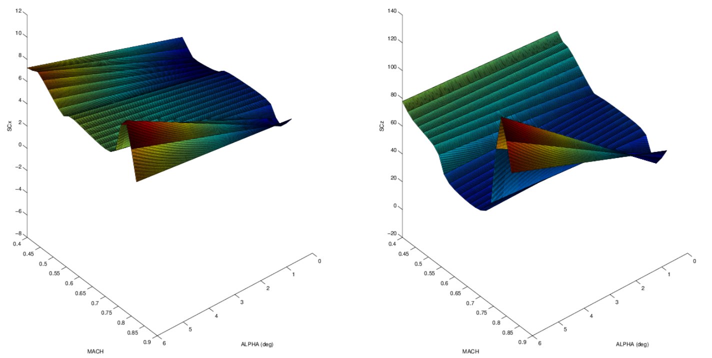


Figure 1. Lift and drag aerodynamic forces for a Boeing 737.

8.1.2. IFPEN

This study is presently conducted in the framework of the PhD of Arthur Le Rhun, started in Fall 2016. The main axis is to design a traffic model suitable for optimizing the fuel consumption of a hybrid vehicle following a given route. The first step was to develop a new traffic model in which the consumption is infered only on the functionning points in the (speed,torque) plane. More precisely, we are interested in the probability distribution of these functionning points when considering a space/time subdivision into road segments and timeframes (see Fig.4). In order to reduce the huge number of distributions obtained, we perform a clustering step using k-means (Fig.5). Since the objects to be clustered are distributions, we choose to use the Wasserstein distance based on optimal transport. The task of computing these Wasserstein barycenters was done by Sinkhorn iterations, and we also developped a variant of stochastic gradient that scales better for huge data sets.

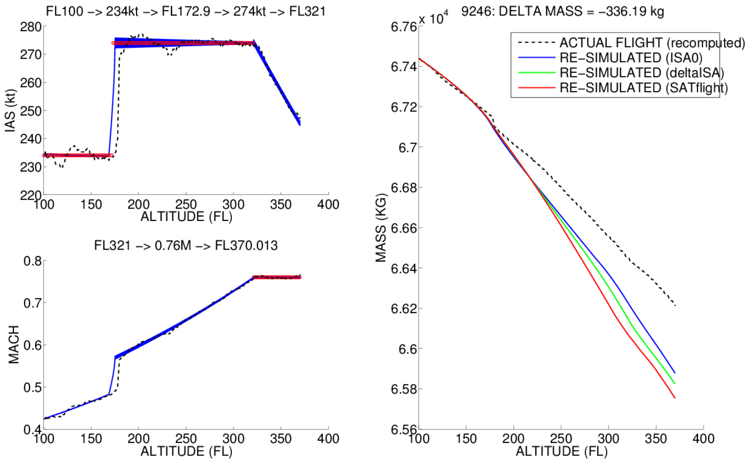


Figure 2. Simulated climb phase vs actual flight data

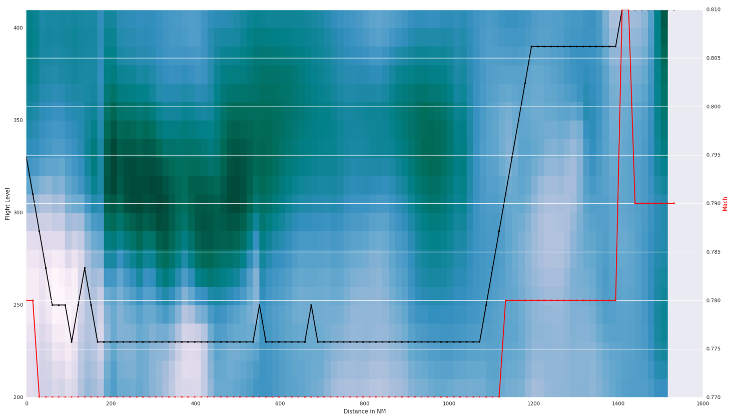


Figure 3. Simulated cruise flight (altitude in black, mach speed in red, wind speed in background)

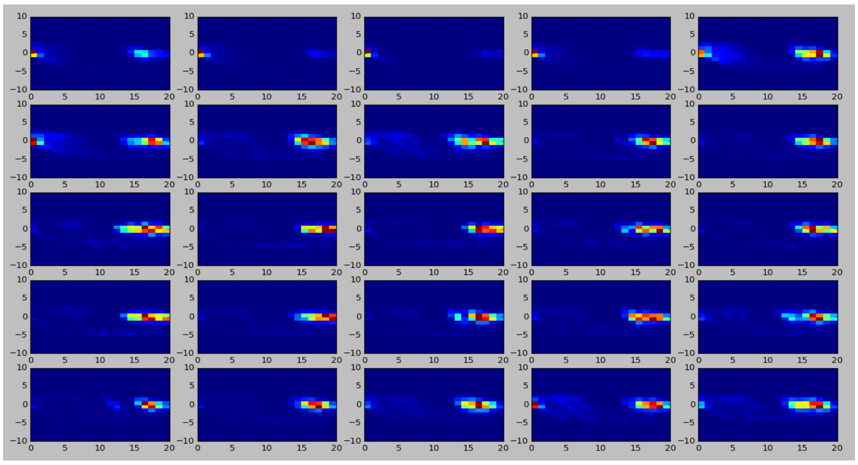


Figure 4. Distributions for all timeframes for a given road segment

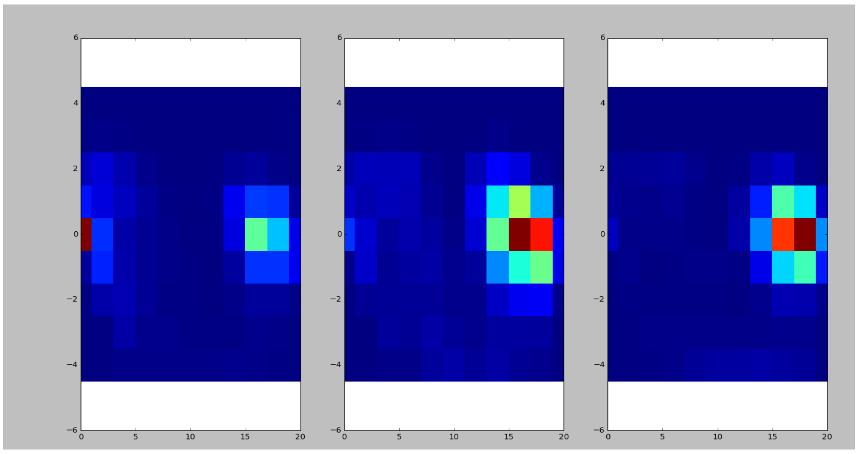


Figure 5. Barycenters after clustering (k=3)

In order to obtain the data for our traffic analysis, we work with a traffic simulator called SUMO, with the LUST scenario modeling the city of Luxembourg (http://sumo.dlr.de).

9. Partnerships and Cooperations

9.1. Regional Initiatives

- Gaspard Monge Program for Optimization and Operational Research (Fondation Jacques Hadamard)

Title	:	Optimal control of partial differential equations using parameterizing manifolds, model reduction, and dynamic programming,
Funding	:	10,000 Euro (for 2016-17), 7,000 Euro (for 2017-2018)
PI	:	Axel Kröner, U. Humboldt and Inria
Period	:	2015 – 2018
Members	:	Frédéric Bonnans (Inria Saclay and CMAP, École Polytechnique), Mickaël Chekroun (UCLA, Los Angeles), Martin Gubisch (U. of Konstanz), Honghu Liu (Virginia Tech), Karl Kunisch (University of Graz), Hasnaa Zidani (ENSTA ParisTech).

9.2. National Initiatives

9.2.1. IPL

9.2.1.1. Cosy

Inria Project Lab COSY (started in 2017) aims at exploiting the potential of state-of-art biological modelling, control techniques, synthetic biology and experimental equipment to achieve a paradigm shift in control of microbial communities. More precisely, we plan to determine and implement control strategies to make heterogeneous communities diversify and interact in the most profitable manner. Study of yeast cells has started in collaboration with team Lifeware (G. Batt) in the framework of the PhD of V. Andreani.

9.2.1.2. Algae in Silico

Inria Project Lab ALGAE IN SILICO (started in 2014) is dedicated to provide an integrated platform for numerical simulation of microalgae "from genes to industrial process". The project has now reached a stage where we can tackle the optimization aspects. Commands is currently joining the IPL, in the following of our previous collaborations with teams Modemic and Biocore on bioreactors, see [35], [23]

9.3. International Research Visitors

9.3.1. Internships

Joao Miguel Machado, from FGV (Rio de Janeiro), spent his master internship in our team from sept-dec 2017, working with F. Bonnans and M.S. Aronna (EMAP-FGV) on the second order necessary and sufficient optimality conditions for optimal control problems of ODEs with broken extremals, i.e., with discontinuous control. We are currently extending the classical theory to the case of a jump between interior and boundary values for the control.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Selection

10.1.1.1. Member of the Conference Program Committees

- F. Bonnans: PGMO Days 2017.

10.1.2. Journal

10.1.2.1. Member of the Editorial Boards

- F. Bonnans: Associate Editor of "Applied Mathematics and Optimization" and of "Series on Mathematics and its Applications, Annals of The Academy of Romanian Scientists".

10.1.2.2. Reviewer - Reviewing Activities

Reviews in 2017 for major journals in the field: Applied Mathematics and Optimization, Automatica, Int. J. of Control, Inverse problems, J. Convex Analysis, J. Diff. Equations, J. of Optimization Theory and Applications, Optimization Set Valued and Variational Analysis, SIAM J. Optimization, SIAM J. Control and Optimization, several conference proceedings.

10.1.3. Invited Talks

- F. Bonnans: Forecasting and risk management for renewable energy, June 7-9, Paris; Numoc, June 19-23, Roma; NHOC2017, July 3-5, Porto; Optimal Control of Partial Differential Equations, Sept, Castro Urdiales.

10.1.4. Leadership within the Scientific Community

- F. Bonnans: French representative to the IFIP-TC7 committee (International Federation of Information Processing; TC7 devoted to System Modeling and Optimization).
- F. Bonnans: member of the PGMO board and Steering Committee (Gaspard Monge Program for Optimization and Operations Research, EDF-FMJH).

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master :
F. Bonnans: *Numerical analysis of partial differential equations arising in finance and stochastic control*, 24h, M2, Ecole Polytechnique and U. Paris 6, France.
F. Bonnans: *Optimal control*, 15h, M2, Optimization master (U. Paris-Saclay) and Ensta, France.
F. Bonnans: *Stochastic optimization*, 15h, M2, Optimization master (U. Paris-Saclay), France.
A. Kröner : Optimal control of partial differential equations, 20h, M2, Optimization master (U. Paris-Saclay), France.

10.2.2. Supervision

PhD in progress : Cédric Rommel, Data exploration for the optimization of aircraft trajectories. Started November 2015 (CIFRE fellowship with Safety Line), F. Bonnans and P. Martinon.

PhD in progress : Arthur Le Rhun, Optimal and robust control of hybrid vehicles. Started September 2016 (IFPEN fellowship), F. Bonnans and P. Martinon.

10.3. Popularization

The collaboration with startup Safety Line was presented at events "Vivatech" (17/06/2017, https://vivatechnology.com/) and "Rencontre Inria Industrie" (17/10/2017) in Paris.

11. Bibliography

Major publications by the team in recent years

[1] A. AFTALION, J. F. BONNANS.*Optimization of running strategies based on anaerobic energy and variations of velocity*, in "SIAM Journal of Applied Mathematics", October 2014, vol. 74, n^{o} 5, p. 1615-1636, https://hal.inria.fr/hal-00851182.

[2] M. S. ARONNA, J. F. BONNANS, B. S. GOH.*Second order analysis of state-constrained control-affine problems*, in "Mathematical Programming, Series A", January 2016, vol. 160, n^{o} 1, p. 115-147, https://hal.inria.fr/hal-01081111.

[3] M. S. ARONNA, J. F. BONNANS, A. KRÖNER.*Optimal Control of Infinite Dimensional Bilinear Systems: Application to the Heat and Wave Equations*, in "Mathematical Programming B", November 2016, 32, https://hal.inria.fr/hal-01273496.

[4] M. S. ARONNA, J. F. BONNANS, P. MARTINON.*A Shooting Algorithm for Optimal Control Problems with Singular Arcs*, in "Journal of Optimization Theory and Applications", August 2013, vol. 158, n^{o} 2, p. 419-459 [*DOI :* 10.1007/S10957-012-0254-8], https://hal.inria.fr/inria-00631332.

[5] J. F. BONNANS, X. DUPUIS, L. PFEIFFER.*Second-order necessary conditions in Pontryagin form for optimal control problems*, in "SIAM Journal on Control and Optimization", 2014, vol. 52, n^{o} 6, p. 3887-3916 [*DOI :* 10.1137/130923452], https://hal.inria.fr/hal-00825273.

[6] J. F. BONNANS, X. TAN.*A model-free no-arbitrage price bound for variance options*, in "Applied Mathematics and Optimization", July 2013, vol. 68, n^{o} 1, p. 43-73 [*DOI :* 10.1007/S00245-013-9197-1], https://hal.inria.fr/inria-00634387.

[7] B. BONNARD, M. CLAEYS, O. COTS, P. MARTINON.*Geometric and numerical methods in the contrast imaging problem in nuclear magnetic resonance*, in "Acta Applicandae Mathematicae", February 2015, vol. 135, n^{o} 1, p. 5-45 [*DOI :* 10.1007/S10440-014-9947-3], https://hal.inria.fr/hal-00867753.

[8] L. GIRALDI, P. MARTINON, M. ZOPPELLO.*Optimal Design for Purcell Three-link Swimmer*, in "Physical Review", February 2015, vol. 91, n^{o} 2, 023012, https://hal.archives-ouvertes.fr/hal-01098501.

[9] A. KRÖNER, K. KUNISCH, B. VEXLER.*Semismooth Newton methods for optimal control of the wave equation with control constraints*, in "SIAM Journal on Control and Optimization", 2011, vol. 49, n^{o} 2, p. 830–858, http://dx.doi.org/10.1137/090766541.

[10] A. KRÖNER, K. KUNISCH, H. ZIDANI.*Optimal feedback control of undamped wave equations by solving a HJB equation*, in "ESAIM: Control, Optimisation and Calculus of Variations", 2014, vol. 21, n^{o} 2, p. 442 - 464 [*DOI :* 10.1051/COCV/2014033], https://hal.archives-ouvertes.fr/hal-00924089.

Publications of the year

Articles in International Peer-Reviewed Journal

[11] J. F. BONNANS, A. FESTA.*Error estimates for the Euler discretization of an optimal control problem with first-order state constraints*, in "SIAM Journal on Numerical Analysis", December 2017, vol. 55, n^{o} 2, p. 445–471, https://hal.inria.fr/hal-01093229.

[12] M. D. CHEKROUN, A. KRÖNER, H. LIU.*Galerkin approximations of nonlinear optimal control problems in Hilbert spaces* , in "Electronic Journal of Differential Equations", 2017, vol. 2017, n^{o} 189, p. 1-40, https://hal.archives-ouvertes.fr/hal-01501178.

[13] J. GARCKE, A. KRÖNER.*Suboptimal feedback control of PDEs by solving HJB equations on adaptive sparse grids*, in "Journal of Scientific Computing", January 2017, vol. 70, n^{o} 1, p. 1-28, https://hal.archives-ouvertes.fr/hal-01185912.

[14] B. HEYMANN, J. F. BONNANS, P. MARTINON, F. SILVA, F. LANAS, G. JIMENEZ.*Continuous Optimal Control Approaches to Microgrid Energy Management*, in "Energy Systems", 2017, https://hal.inria.fr/hal-01129393.

International Conferences with Proceedings

[15] J. F. BONNANS, B. GREGORUTTI, P. MARTINON, C. ROMMEL.*Aircraft Dynamics Identification for Optimal Control*, in "EUCASS 2017", Milan, Italy, R. MARTINEZ-VAL, C. BONNAL, D. KNIGHT (editors), July 2017 [*DOI :* 10.13009/EUCASS2017-179], https://hal.inria.fr/hal-01639731.

Research Reports

[16] M. D. CHEKROUN, A. KRÖNER, H. LIU.*Galerkin approximations for the optimal control of nonlinear delay differential equations*, Inria, 2017, https://hal.archives-ouvertes.fr/hal-01534673.

Other Publications

[17] J. F. BONNANS, J. GIANATTI, F. J. SILVA.*On the time discretization of stochastic optimal control problems: the dynamic programming approach*, February 2017, working paper or preprint, https://hal.inria.fr/hal-01474285.

[18] J. F. BONNANS, A. KRÖNER.*Variational analysis for options with stochastic volatility and multiple factors*, April 2017, working paper or preprint, https://hal.inria.fr/hal-01516011.

[19] A. KRÖNER, A. PICARELLI, H. ZIDANI.*Infinite Horizon Stochastic Optimal Control Problems with Running Maximum Cost*, October 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01585766.

[20] P. MARTINON, P. PICARD, A. RAJ.*On the Design of Optimal Health Insurance Contracts under Ex Post Moral Hazard*, 2017, working paper or preprint, https://hal-polytechnique.archives-ouvertes.fr/hal-01348551.

References in notes

[21] M. BARDI, I. CAPUZZO-DOLCETTA.*Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, Systems and Control: Foundations and Applications, Birkhäuser, Boston, 1997.

[22] G. BARLES.*Solutions de viscosité des équations de Hamilton-Jacobi*, Mathématiques et Applications, Springer, Paris, 1994, vol. 17.

[23] T. BAYEN, F. MAIRET, P. MARTINON, M. SEBBAH.*Analysis of a periodic optimal control problem connected to microalgae anaerobic digestion*, in "Optimal Control Applications and Methods", 2014 [*DOI :* 10.1002/OCA.2127], https://hal.archives-ouvertes.fr/hal-00860570.

[24] G. BLISS.*Lectures on the Calculus of Variations*, University of Chicago Press, Chicago, Illinois, 1946.

[25] J. F. BONNANS, J. LAURENT-VARIN.*Computation of order conditions for symplectic partitioned Runge-Kutta schemes with application to optimal control*, in "Numerische Mathematik", 2006, vol. 103, n^o 1, p. 1–10.

[26] J. F. BONNANS, E. OTTENWAELTER, H. ZIDANI.*Numerical schemes for the two dimensional second-order HJB equation*, in "ESAIM: M2AN", 2004, vol. 38, p. 723-735.

[27] J. F. BONNANS, H. ZIDANI.*Consistency of generalized finite difference schemes for the stochastic HJB equation*, in "SIAM J. Numerical Analysis", 2003, vol. 41, p. 1008-1021.

[28] A. E. BRYSON, Y.-C. HO.*Applied optimal control*, Hemisphere Publishing, New-York, 1975.

[29] M. CRANDALL, P. LIONS.*Viscosity solutions of Hamilton Jacobi equations*, in "Bull. American Mathematical Society", 1983, vol. 277, p. 1–42.

[30] W. HAGER.*Runge-Kutta methods in optimal control and the transformed adjoint system*, in "Numerische Mathematik", 2000, vol. 87, n^o 2, p. 247–282.

[31] A. IOFFE, V. TIHOMIROV.*Theory of Extremal Problems*, North-Holland Publishing Company, Amsterdam, 1979, Russian Edition: Nauka, Moscow, 1974.

[32] J.-L. LAGRANGE.*Mécanique analytique*, Paris, 1788, reprinted by J. Gabay, 1989.

[33] E. LEE, L. MARKUS.*Foundations of optimal control theory*, John Wiley, New York, 1967.

[34] G. LEITMANN.*An introduction to optimal control*, Mc Graw Hill, New York, 1966.

[35] R. MUÑOZ-TAMAYO, P. MARTINON, G. BOUGARAN, F. MAIRET, O. BERNARD.*Getting the most out of it: Optimal experiments for parameter estimation of microalgae growth models*, in "Journal of Process Control", 2014, vol. 24, n^o 6, p. 991 - 1001 [*DOI :* 10.1016/J.JPROCONT.2014.04.021], http://www.sciencedirect.com/science/article/pii/S095915241400122X.

[36] L. PONTRYAGIN, V. BOLTYANSKI, R. GAMKRELIDZE, E. MICHTCHENKO.*The Mathematical Theory of Optimal Processes*, Wiley Interscience, New York, 1962.

[37] L. YOUNG.*Lectures on the calculus of variations and optimal control theory*, W. B. Saunders Co., Philadelphia, 1969.

Project-Team DATASHAPE

Understanding the shape of data

RESEARCH CENTERS
Saclay - Île-de-France
Sophia Antipolis - Méditerranée

THEME
Algorithmics, Computer Algebra and Cryptology

Table of contents

Project-Team DATASHAPE

Creation of the Team: 2016 January 01, updated into Project-Team: 2016 January 01

Keywords:

Computer Science and Digital Science:

A3. - Data and knowledge
A3.4. - Machine learning and statistics
A7.1. - Algorithms
A8. - Mathematics of computing
A8.1. - Discrete mathematics, combinatorics
A8.3. - Geometry, Topology
A9. - Artificial intelligence

Other Research Topics and Application Domains:

B1. - Life sciences
B2. - Health
B5. - Industry of the future
B9. - Society and Knowledge
B9.4. - Sciences

1. Personnel

Research Scientists

Jean-Daniel Boissonnat [Inria, Senior Researcher, HDR]
Frédéric Chazal [Team leader, Inria, Senior Researcher, HDR]
David Cohen-Steiner [Inria, Researcher]
Kunal Dutta [Inria, Starting Research Position, from Oct 2017]
Marc Glisse [Inria, Researcher]
Miroslav Kramar [Inria, Advanced Research Position, from Apr 2017]
Steve Oudot [Inria, Researcher, HDR]
Mathijs Wintraecken [Inria, Starting Research Position, from Nov 2017]

Faculty Members

Guilherme Dias Da Fonseca [Univ d'Auvergne Clermont-Ferrand, Associate Professor, from Sep 2017]
Maksims Ovsjanikovs [Ecole Polytechnique, Professor]

External Collaborators

Eddie Aamari [UC San Diego, from Sep 2017]
Pawel Dlotko [University of Swansea, from Mar 2017]
Ramsay Dyer [Mathematical Sciences Publishers]
Clément Maria [University of Brisbane, until Nov 2017]
Konstantin Mischaikow [Rutgers University, Nov 2017]

Technical Staff

Clément Jamin [Inria, until Nov 2017]
Vincent Rouvreau [Inria, Engineer SED 40%]

PhD Students

Eddie Aamari [Région, until Aug 2017]
Bertrand Beaufils [Sysnav, industrial grant]

Maxime Bellec [Sysnav, industrial grant, from Oct 2017]
Nicolas Berkouk [Inria]
Claire Brecheteau [Univ Paris-Sud]
Mathieu Carrière [Inria]
Jeremy Sylvain Cochoy [Univ Paris-Sud]
Alba Chiara de Vitis [Inria, until Jul 2017]
Francois Godi [Univ de Nice - Sophia Antipolis]
Siargey Kachanovich [Inria]
Theo Lacombe [Ecole Polytechnique, from Sep 2017]
Siddharth Pritam [Inria]
Raphael Tinarrage [Ecole Normale Supérieure Cachan, from Sep 2017]

Post-Doctoral Fellows

Kunal Dutta [Inria, until Sep 2017]
Mathijs Wintraecken [Inria, until Oct 2017]

Visiting Scientists

Magnus Bakke Botnan [TU München, Mar 2017]
Arijit Ghosh [Indian Statistical Institute, June 2017 and November 2017]
Ilaria Giulini [Univ Pierre et Marie Curie]
Bruno Jartoux [ESIEE, Jul 2017]
Manish Krishan Lal [IIT Madras, from May 2017 until Jul 2017]
Theo Lacombe [Ecole Polytechnique, from Apr 2017 until Aug 2017]
Bertrand Michel [Univ Pierre et Marie Curie]
Mariette Yvinec [Retraite, HDR]

Administrative Assistants

Christine Aklouche [Inria, until Aug 2017]
Florence Barbara [Inria, Assistant]

2. Overall Objectives

2.1. Overall Objectives

DataShape is a research project in Topological Data Analysis (TDA), a recent field whose aim is to uncover, understand and exploit the topological and geometric structure underlying complex and possibly high dimensional data. The DATASHAPE project gathers a unique variety of expertise that allows it to embrace the mathematical, statistical, algorithmic and applied aspects of the field in a common framework ranging from fundamental theoretical studies to experimental research and software development.

The expected output of DATASHAPE is two-fold. First, we intend to set-up and develop the mathematical, statistical and algorithmic foundations of Topological and Geometric Data Analysis. Second, we intend to develop the Gudhi platform in order to provide an efficient state-of-the-art toolbox for the understanding of the topology and geometry of data.

3. Research Program

3.1. Algorithmic aspects of topological and geometric data analysis

TDA requires to construct and manipulate appropriate representations of complex and high dimensional shapes. A major difficulty comes from the fact that the complexity of data structures and algorithms used to approximate shapes rapidly grows as the dimensionality increases, which makes them intractable in high dimensions. We focus our research on simplicial complexes which offer a convenient representation of general shapes and generalize graphs and triangulations. Our work includes the study of simplicial complexes with good approximation properties and the design of compact data structures to represent them.

In low dimensions, effective shape reconstruction techniques exist that can provide precise geometric approximations very efficiently and under reasonable sampling conditions. Extending those techniques to higher dimensions as is required in the context of TDA is problematic since almost all methods in low dimensions rely on the computation of a subdivision of the ambient space. A direct extension of those methods would immediately lead to algorithms whose complexities depend exponentially on the ambient dimension, which is prohibitive in most applications. A first direction to by-pass the curse of dimensionality is to develop algorithms whose complexities depend on the intrinsic dimension of the data (which most of the time is small although unknown) rather than on the dimension of the ambient space. Another direction is to resort to cruder approximations that only captures the homotopy type or the homology of the sampled shape. The recent theory of persistent homology provides a powerful and robust tool to study the homology of sampled spaces in a stable way.

3.2. Statistical aspects of topological and geometric data analysis

The wide variety of larger and larger available data - often corrupted by noise and outliers - requires to consider the statistical properties of their topological and geometric features and to propose new relevant statistical models for their study.

There exist various statistical and machine learning methods intending to uncover the geometric structure of data. Beyond manifold learning and dimensionality reduction approaches that generally do not allow to assert the relevance of the inferred topological and geometric features and are not well-suited for the analysis of complex topological structures, set estimation methods intend to estimate, from random samples, a set around which the data is concentrated. In these methods, that include support and manifold estimation, principal curves/manifolds and their various generalizations to name a few, the estimation problems are usually considered under losses, such as Hausdorff distance or symmetric difference, that are not sensitive to the topology of the estimated sets, preventing these tools to directly infer topological or geometric information.

Regarding purely topological features, the statistical estimation of homology or homotopy type of compact subsets of Euclidean spaces, has only been considered recently, most of the time under the quite restrictive assumption that the data are randomly sampled from smooth manifolds.

In a more general setting, with the emergence of new geometric inference tools based on the study of distance functions and algebraic topology tools such as persistent homology, computational topology has recently seen an important development offering a new set of methods to infer relevant topological and geometric features of data sampled in general metric spaces. The use of these tools remains widely heuristic and until recently there were only a few preliminary results establishing connections between geometric inference, persistent homology and statistics. However, this direction has attracted a lot of attention over the last three years. In particular, stability properties and new representations of persistent homology information have led to very promising results to which the DATASHAPE members have significantly contributed. These preliminary results open many perspectives and research directions that need to be explored.

Our goal is to build on our first statistical results in TDA to develop the mathematical foundations of Statistical Topological and Geometric Data Analysis. Combined with the other objectives, our ultimate goal is to provide a well-founded and effective statistical toolbox for the understanding of topology and geometry of data.

3.3. Topological approach for multimodal data processing

Due to their geometric nature, multimodal data (images, video, 3D shapes, etc.) are of particular interest for the techniques we develop. Our goal is to establish a rigorous framework in which data having different representations can all be processed, mapped and exploited jointly. This requires adapting our tools and sometimes developing entirely new or specialized approaches.

The choice of multimedia data is motivated primarily by the fact that the amount of such data is steadily growing (with e.g. video streaming accounting for nearly two thirds of peak North-American Internet traffic, and almost half a billion images being posted on social networks each day), while at the same time it poses significant challenges in designing informative notions of (dis)-similarity as standard metrics (e.g. Euclidean distances between points) are not relevant.

3.4. Experimental research and software development

We develop a high quality open source software platform called GUDHI which is becoming a reference in geometric and topological data analysis in high dimensions. The goal is not to provide code tailored to the numerous potential applications but rather to provide the central data structures and algorithms that underly applications in geometric and topological data analysis.

The development of the GUDHI platform also serves to benchmark and optimize new algorithmic solutions resulting from our theoretical work. Such development necessitates a whole line of research on software architecture and interface design, heuristics and fine-tuning optimization, robustness and arithmetic issues, and visualization. We aim at providing a full programming environment following the same recipes that made up the success story of the CGAL library, the reference library in computational geometry.

Some of the algorithms implemented on the platform will also be interfaced to other software platform, such as the R software [0] for statistical computing, and languages such as Python in order to make them usable in combination with other data analysis and machine learning tools. A first attempt in this direction has been done with the creation of an R package called TDA in collaboration with the group of Larry Wasserman at Carnegie Mellon University (Inria Associated team CATS) that already includes some functionalities of the GUDHI library and implements some joint results between our team and the CMU team. A similar interface with the Python language is also considered a priority. To go even further towards helping users, we will provide utilities that perform the most common tasks without requiring any programming at all.

4. Application Domains

4.1. Main application domains

Our work is mostly of a fundamental mathematical and algorithmic nature but fin ds applications in a variety of application in data analysis, more precisely in Topological Data Analysis (TDA). Although TDA is a quite recent field, it al ready founds applications in material science, biology, sensor networks, 3D shap es analysis and processing, to name a few.

More specifically, DATASHAPEhas recently started to work on the analysis of t rajectories obtained from inertial sensors (starting PhD thesis of Bertrand Bea ufils) and is exploring some possible new applications in material science.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Chairs

Jean-Daniel Boissonnat was elected a professor at the Collège de France, on the Chair Informatics and Computational Sciences for the academic year 2016-2017.

6. New Software and Platforms

6.1. GUDHI

Geometric Understanding in Higher Dimensions
KEYWORDS: Computational geometry - Topology

[0] https://www.r-project.org/

SCIENTIFIC DESCRIPTION: The current release of the GUDHI library includes: – Data structures to represent, construct and manipulate simplicial and cubical complexes. – Algorithms to compute simplicial complexes from point cloud data. – Algorithms to compute persistent homology and multi-field persistent homology. – Simplification methods via implicit representations.
FUNCTIONAL DESCRIPTION: The GUDHI open source library will provide the central data structures and algorithms that underly applications in geometry understanding in higher dimensions. It is intended to both help the development of new algorithmic solutions inside and outside the project, and to facilitate the transfer of results in applied fields.

RELEASE FUNCTIONAL DESCRIPTION: Major new features in 2017: - python interface - bottleneck distance - tangential complex - relaxed witness complex

- Participants: Clément Maria, François Godi, David Salinas, Jean-Daniel Boissonnat, Marc Glisse, Mariette Yvinec, Pawel Dlotko, Siargey Kachanovich and Vincent Rouvreau
- Contact: Jean-Daniel Boissonnat
- URL: http://gudhi.gforge.inria.fr/

6.2. dD Triangulations

CGAL module: Triangulations in any dimension
KEYWORDS: 3D modeling - Triangulation - Delaunay triangulation - Voronoi diagram - Regular triangulation
FUNCTIONAL DESCRIPTION: This package of CGAL (Computational Geometry Algorithms Library http://www.cgal.org) allows to compute triangulations, Delaunay triangulations and regular triangulations in any dimension. Those triangulations are built incrementally and can be modified by insertion or removal of vertices.

RELEASE FUNCTIONAL DESCRIPTION: Version 4.11 adds the regular triangulations to the package.

- Participants: Clément Jamin, Olivier Devillers and Samuel Hornus
- Contact: Samuel Hornus
- URL: http://www.cgal.org

7. New Results

7.1. Algorithmic aspects of topological and geometric data analysis

7.1.1. Variance Minimizing Transport Plans for Inter-surface Mapping

Participant: David Cohen-Steiner.

In collaboration with Manish Mandad, Leik Kobbelt (RWTH Aachen), Pierre Alliez (Inria), and Mathieu Desbrun (Caltech).

We introduce an effcient computational method for generating dense and low distortion maps between two arbitrary surfaces of same genus. Instead of relying on semantic correspondences or surface parameterization, we directly optimize a variance-minimizing transport plan between two input surfaces that defines an as-conformal-as-possible inter-surface map satisfying a user-prescribed bound on area distortion. The transport plan is computed via two alternating convex optimizations, and is shown to minimize a generalized Dirichlet energy of both the map and its inverse. Computational efficiency is achieved through a coarse-tone approach in diffusion geometry, with Sinkhorn iterations modified to enforce bounded area distortion. The resulting inter-surface mapping algorithm applies to arbitrary shapes robustly, with little to no user interaction.

7.1.2. Approximating the spectrum of a graph

Participant: David Cohen-Steiner.

In collaboration with Weihao Kong, Gregory Valiant (Stanford), and Christian Sohler (TU Dortmund).

The spectrum of a network or graph $G = (V, E)$ with adjacency matrix A consists of the eigenvalues of the normalized Laplacian $L = I - D^{-1/2}AD^{-1/2}$. This set of eigenvalues encapsulates many aspects of the structure of the graph, including the extent to which the graph posses community structures at multiple scales. We study the problem of approximating the spectrum $\lambda = (\lambda_1, \ldots, \lambda_{|V|})$, $0 \leq \lambda 1, \leq \ldots, \leq \lambda |V| \leq 2$ of G in the regime where the graph is too large to explicitly calculate the spectrum. We present a sublinear time algorithm that, given the ability to query a random node in the graph and select a random neighbor of a given node, computes a succinct representation of an approximation $\widetilde{\lambda}$ such that $\|\widetilde{\lambda} - \lambda\|_1 \leq \varepsilon|V|$. Our algorithm has query complexity and running time $\exp(O(1/\varepsilon))$, independent of the size of the graph, $|V|$. We demonstrate the practical viability of our algorithm on 15 different real-world graphs from the Stanford Large Network Dataset Collection, including social networks, academic collaboration graphs, and road networks. For the smallest of these graphs, we are able to validate the accuracy of our algorithm by explicitly calculating the true spectrum; for the larger graphs, such a calculation is computationally prohibitive. In addition we study the implications of our algorithm to property testing in the bounded degree graph model.

7.1.3. Anisotropic triangulations via discrete Riemannian Voronoi diagrams

Participants: Jean-Daniel Boissonnat, Mathijs Wintraecken.

In collaboration with mael Rouxel-Labbé (GeometryFactory).

The construction of anisotropic triangulations is desirable for various applications, such as the numerical solving of partial differential equations and the representation of surfaces in graphics. To solve this notoriously difficult problem in a practical way, we introduce the discrete Riemannian Voronoi diagram, a discrete structure that approximates the Riemannian Voronoi diagram. This structure has been implemented and was shown to lead to good triangulations in $\mathbb{R}^2$ and on surfaces embedded in $\mathbb{R}^3$ as detailed in our experimental companion paper.

In [23], [32], [34], we study theoretical aspects of our structure. Given a finite set of points $\mathcal{P}$ in a domain Ω equipped with a Riemannian metric, we compare the discrete Riemannian Voronoi diagram of $\mathcal{P}$ to its Riemannian Voronoi diagram. Both diagrams have dual structures called the discrete Riemannian Delaunay and the Riemannian Delaunay complex. We provide conditions that guarantee that these dual structures are identical. It then follows from previous results that the discrete Riemannian Delaunay complex can be embedded in Ω under sufficient conditions, leading to an anisotropic triangulation with curved simplices. Furthermore, we show that, under similar conditions, the simplices of this triangulation can be straightened.

7.1.4. Only distances are required to reconstruct submanifolds

Participants: Jean-Daniel Boissonnat, Ramsay Dyer, Steve Oudot.

In collaboration with Arijit Ghosh (Indian Statistical Institute).

In [14], we give the first algorithm that outputs a faithful reconstruction of a submanifold of Euclidean space without maintaining or even constructing complicated data structures such as Voronoi diagrams or Delaunay complexes. Our algorithm uses the witness complex and relies on the stability of *power protection*, a notion introduced in this paper. The complexity of the algorithm depends exponentially on the intrinsic dimension of the manifold, rather than the dimension of ambient space, and linearly on the dimension of the ambient space. Another interesting feature of this work is that no explicit coordinates of the points in the point sample is needed. The algorithm only needs the *distance matrix* as input, i.e., only distance between points in the point sample as input.

7.1.5. An obstruction to Delaunay triangulations in Riemannian manifolds

Participants: Jean-Daniel Boissonnat, Ramsay Dyer.

In collaboration with Arijit Ghosh (Indian Statistical Institute) and Nikolay Martynchuk (University of Groningen).

Delaunay has shown that the Delaunay complex of a finite set of points P of Euclidean space $\mathbb{R}^m$ triangulates the convex hull of P, provided that P satisfies a mild genericity property. Voronoi diagrams and Delaunay complexes can be defined for arbitrary Riemannian manifolds. However, Delaunay's genericity assumption no longer guarantees that the Delaunay complex will yield a triangulation; stronger assumptions on P are required. A natural one is to assume that P is sufficiently dense. Although results in this direction have been claimed, we show that sample density alone is insufficient to ensure that the Delaunay complex triangulates a manifold of dimension greater than 2 [13].

7.1.6. Local criteria for triangulation of manifolds

Participants: Jean-Daniel Boissonnat, Ramsay Dyer, Mathijs Wintraecken.

In collaboration with Arijit Ghosh (Indian Statistical Institute).

We present criteria for establishing a triangulation of a manifold [40]. Given a manifold M, a simplicial complex $\mathcal{A}$, and a map H from the underlying space of $\mathcal{A}$ to M, our criteria are presented in local coordinate charts for M, and ensure that H is a homeomorphism. These criteria do not require a differentiable structure, or even an explicit metric on M. No Delaunay property of $\mathcal{A}$ is assumed. The result provides a triangulation guarantee for algorithms that construct a simplicial complex by working in local coordinate patches. Because the criteria are easily checked algorithmically, they are expected to be of general use.

7.1.7. Triangulating stratified manifolds I: a reach comparison theorem

Participants: Jean-Daniel Boissonnat, Mathijs Wintraecken.

In [42], we define the reach for submanifolds of Riemannian manifolds, in a way that is similar to the Euclidean case. Given a d-dimensional submanifold $\mathcal{S}$ of a smooth Riemannian manifold $\mathbb{M}$ and a point $p \in \mathbb{M}$ that is not too far from $\mathcal{S}$ we want to give bounds on local feature size of $\exp_p^{-1}(\mathcal{S})$. Here $\exp_p^{-1}$ is the inverse exponential map, a canonical map from the manifold to the tangent space. Bounds on the local feature size of $\exp_p^{-1}(\mathcal{S})$ can be reduced to giving bounds on the reach of $\exp_p^{-1}(\mathcal{B})$, where $\mathcal{B}$ is a geodesic ball, centred at c with radius equal to the reach of $\mathcal{S}$. Equivalently we can give bounds on the reach of $\exp_p^{-1} \circ \exp_c(B_c)$, where now B_c is a ball in the tangent space $T_c\mathbb{M}$, with the same radius. To establish bounds on the reach of $\exp_p^{-1} \circ \exp_c(B_c)$ we use bounds on the metric and on its derivative in Riemann normal coordinates.

This result is a first step towards answering the important question of how to triangulate stratified manifolds.

7.1.8. The reach, metric distortion, geodesic convexity and the variation of tangent spaces

Participants: Jean-Daniel Boissonnat, Mathijs Wintraecken.

In collaboration with André Lieutier (Dassault Système).

In [41], we discuss three results. The first two concern general sets of positive reach: We first characterize the reach by means of a bound on the metric distortion between the distance in the ambient Euclidean space and the set of positive reach. Secondly, we prove that the intersection of a ball with radius less than the reach with the set is geodesically convex, meaning that the shortest path between any two points in the intersection lies itself in the intersection. For our third result we focus on manifolds with positive reach and give a bound on the angle between tangent spaces at two different points in terms of the distance between the points and the reach.

7.1.9. Delaunay triangulation of a random sample of a good sample has linear size

Participants: Jean-Daniel Boissonnat, Kunal Dutta, Marc Glisse.

In collaboration with Olivier Devillers (Inria Nancy Grand Est).

The *randomized incremental construction* (RIC) for building geometric data structures has been analyzed extensively, from the point of view of worst-case distributions. In many practical situations however, we have to face nicer distributions. A natural question that arises is: do the usual RIC algorithms automatically adapt when the point samples are nicely distributed. We answer positively to this question for the case of the Delaunay triangulation of ϵ-nets.

ϵ-nets are a class of nice distributions in which the point set is such that any ball of radius ϵ contains at least one point of the net and two points of the net are distance at least ϵ apart. The Delaunay triangulations of ϵ-nets are proved to have linear size; unfortunately this is not enough to ensure a good time complexity of the randomized incremental construction of the Delaunay triangulation. In [33], [38], we prove that a uniform random sample of a given size that is taken from an ϵ-net has a linear sized Delaunay triangulation in any dimension. This result allows us to prove that the randomized incremental construction needs an expected linear size and an expected $O(n \log n)$ time.

Further, we also prove similar results in the case of non-Euclidean metrics, when the point distribution satisfies a certain *bounded expansion* property; such metrics can occur, for example, when the points are distributed on a low-dimensional manifold in a high-dimensional ambient space.

7.1.10. Kernelization of the Subset General Position problem in Geometry

Participants: Jean-Daniel Boissonnat, Kunal Dutta.

In collaboration with Arijit Ghosh (Indian Statistical Institute) and Sudeshna Kolay (Eindhoven University of Technology).

In [21], we consider variants of the GEOMETRIC SUBSET GENERAL POSITION problem. In defining this problem, a geometric subsystem is specified, like a subsystem of lines, hyperplanes or spheres. The input of the problem is a set of n points in $\mathbb{R}^d$ and a positive integer k. The objective is to find a subset of at least k input points such that this subset is in general position with respect to the specified subsystem. For example, a set of points is in general position with respect to a subsystem of hyperplanes in $\mathbb{R}^d$ if no $d+1$ points lie on the same hyperplane. In this paper, we study the HYPERPLANE SUBSET GENERAL POSITION problem under two parameterizations. When parameterized by k then we exhibit a polynomial kernelization for the problem. When parameterized by $h = n - k$, or the dual parameter, then we exhibit polynomial kernels which are also tight, under standard complexity theoretic assumptions. We can also conclude similar kernelization results for D-POLYNOMIAL SUBSET GENERAL POSITION, where a vector space of polynomials of degree at most d are specified as the underlying subsystem such that the size of the basis for this vector space is b. The objective is to find a set of at least k input points, or in the dual delete at most $h = n - k$ points, such that no $b+1$ points lie on the same polynomial. Notice that this is a generalization of many well-studied geometric variants of the SET COVER problem, such as CIRCLE SUBSET GENERAL POSITION. We also study the general projective variants of these problems. These problems are also related to other geometric problems like SUBSET DELAUNAY TRIANGULATION problem.

7.1.11. Tight Kernels for Covering and Hitting: Point Hyperplane Cover and Polynomial Point Hitting Set

Participants: Jean-Daniel Boissonnat, Kunal Dutta.

In collaboration with Arijit Ghosh (Indian Statistical Institute) and Sudeshna Kolay (Eindhoven University of Technology).

The POINT HYPERPLANE COVER problem in $\mathbb{R}^d$ takes as input a set of n points in $\mathbb{R}^d$ and a positive integer k. The objective is to cover all the given points with a set of at most k hyperplanes. The D-POLYNOMIAL POINTS HITTING SET (D-POLYNOMIAL POINTS HS) problem in $\mathbb{R}^d$ takes as input a family $\mathcal{F}$ of D-degree polynomials from a vector space $\mathcal{R}$ in $\mathbb{R}^d$, and determines whether there is a set of at most k points in $\mathbb{R}^d$ that hit all the polynomials in $\mathcal{F}$. In [22], we exhibit tight kernels where k is the parameter for these problems.

7.1.12. Shallow packings, semialgebraic set systems, Macbeath regions, and polynomial partitioning

Participant: Kunal Dutta.

In collaboration with Arijit Ghosh (Indian Statistical Institute) and Bruno Jartoux (Université Paris-Est, Laboratoire d'Informatique Gaspard-Monge, ESIEE Paris, France) and Nabil H. Mustafa (Université Paris-Est, Laboratoire d'Informatique Gaspard-Monge, ESIEE Paris, France).

The packing lemma of Haussler states that given a set system $(X, \mathbb{R})$ with bounded VC dimension, if every pair of sets in $\mathbb{R}$ have large symmetric difference, then $\mathbb{R}$ cannot contain too many sets. Recently it was generalized to the shallow packing lemma, applying to set systems as a function of their shallow-cell complexity. In [29] we present several new results and applications related to packings:

1. an optimal lower bound for shallow packings,
2. improved bounds on Mnets, providing a combinatorial analogue to Macbeath regions in convex geometry,
3. we observe that Mnets provide a general, more powerful framework from which the state-of-the-art unweighted ϵ-net results follow immediately, and
4. simplifying and generalizing one of the main technical tools in Fox *et al.* (*J. of the EMS*, to appear).

7.1.13. A Simple Proof of Optimal Epsilon Nets

Participant: Kunal Dutta.

In collaboration with Nabil H. Mustafa (Université Paris-Est, Laboratoire d'Informatique Gaspard-Monge, ESIEE Paris, France, and Arijit Ghosh (Indian Statistical Institute)).

Showing the existence of ϵ-nets of small size has been the subject of investigation for almost 30 years, starting from the initial breakthrough of Haussler and Welzl (1987). Following a long line of successive improvements, recent results have settled the question of the size of the smallest ϵ-nets for set systems as a function of their so-called shallow-cell complexity.

In [20] we give a short proof of this theorem in the space of a few elementary paragraphs, showing that it follows by combining the ϵ-net bound of Haussler and Welzl (1987) with a variant of Haussler's packing lemma (1991).

This implies all known cases of results on unweighted ϵ-nets studied for the past 30 years, starting from the result of Matoušek, Seidel and Welzl (1990) to that of Clarkson and Varadajan (2007) to that of Varadarajan (2010) and Chan, Grant, Könemann and Sharpe (2012) for the unweighted case, as well as the technical and intricate paper of Aronov, Ezra and Sharir (2010).

7.1.14. On Subgraphs of Bounded Degeneracy in Hypergraphs

Participant: Kunal Dutta.

In collaboration with Arijit Ghosh (Indian Statistical Institute)).

A k-uniform hypergraph is d-degenerate if every induced subgraph has a vertex of degree at most d. In [48], given a k-uniform hypergraph $H = (V(H), E(H))$, we show there exists an induced subgraph of size at least

$$\sum_{v \in V(H)} \min \left\{ 1,\ c_k \left(\frac{d+1}{d_H(v)+1} \right)^{1/(k-1)} \right\},$$

where $c_k = 2^{-\left(1+\frac{1}{k-1}\right)} \left(1 - \frac{1}{k}\right)$ and $d_H(v)$ denotes the degree of vertex v in the hypergraph H. This connects, extends, and generalizes results of Alon-Kahn-Seymour (1987), on d-degenerate sets of graphs, Dutta-Mubayi-Subramanian (2012) on d-degenerate sets of linear hypergraphs, and Srinivasan-Shachnai (2004) on independent sets in hypergraphs to d-degenerate subgraphs of hypergraphs. Our technique also gives optimal lower bounds for a more generalized definition of degeneracy introduced by Zaker (2013). We further give a simple non-probabilistic proof of the Dutta-Mubayi-Subramanian bound for linear k-uniform hypergraphs, which extends the Alon-Kahn-Seymour proof technique to hypergraphs. Finally we provide several applications in discrete geometry, extending results of Payne-Wood (2013) and Cardinal-Tóth-Wood (2016). We also address some natural algorithmic questions. The proof of our main theorem combines the *random permutation* technique of Bopanna-Caro-Wei and Beame and Luby, together with a new *local density* argument which may be of independent interest.

7.2. Statistical aspects of topological and geometric data analysis

7.2.1. The DTM-signature for a geometric comparison of metric-measure spaces from samples

Participant: Claire Brécheteau.

In [43], we introduce the notion of DTM-signature, a measure on $\mathbb{R}_+$ that can be associated to any metric-measure space. This signature is based on the distance to a measure (DTM) introduced by Chazal, Cohen-Steiner and Mérigot. It leads to a pseudo-metric between metric-measure spaces, upper-bounded by the Gromov-Wasserstein distance. Under some geometric assumptions, we derive lower bounds for this pseudo-metric. Given two N-samples, we also build an asymptotic statistical test based on the DTM-signature, to reject the hypothesis of equality of the two underlying metric measure spaces, up to a measure-preserving isometry. We give strong theoretical justifications for this test and propose an algorithm for its implementation.

7.2.2. Estimating the Reach of a Manifold

Participants: Eddie Aamari, Frédéric Chazal, Bertrand Michel.

In collaboration with J. Kim, A. Rinaldo, L. Wasserman (Carnegie Mellon University)

Various problems of computational geometry and manifold learning encode geometric regularity through the so-called reach, a generalized convexity parameter. The reach τ_M of a submanifold $M \subset \mathbb{R}^D$ is the maximal offset radius on which the projection onto M is well defined. The quantity τ_M renders a certain minimal scale of M, giving bounds on both maximum curvature and possible bottleneck structures. In [35], we study the geometry of the reach through an approximation perspective. We derive new geometric results on the reach for submanifolds without boundary. An estimator $\widehat{\tau}$ of τ_M is proposed in a framework where tangent spaces are known, and bounds assessing its efficiency are derived. In the case of i.i.d. random point cloud $\mathbb{X}_n$, $\widehat{\tau}(\mathbb{X}_n)$ is showed to achieve uniform expected loss bounds over a $\mathcal{C}^3$-like model. Minimax upper and lower bounds are derived, and we conclude with the extension to a model with unknown tangent spaces.

7.2.3. Robust Topological Inference: Distance To a Measure and Kernel Distance

Participants: Frédéric Chazal, Bertrand Michel.

In collaboration with B. Fasy, F. Lecci, A. Rinaldo, L. Wasserman.

Let P be a distribution with support S. The salient features of S can be quantified with persistent homology, which summarizes topological features of the sublevel sets of the distance function (the distance of any point x to S). Given a sample from P we can infer the persistent homology using an empirical version of the distance function. However, the empirical distance function is highly non-robust to noise and outliers. Even one outlier is deadly. The distance-to-a-measure (DTM) and the kernel distance are smooth functions that provide useful topological information but are robust to noise and outliers. In [17], we derive limiting distributions and confidence sets, and we propose a method for choosing tuning parameters.

7.2.4. Statistical analysis and parameter selection for Mapper

Participants: Steve Oudot, Bertrand Michel, Mathieu Carrière.

In [44] we study the question of the statistical convergence of the 1-dimensional Mapper to its continuous analogue, the Reeb graph. We show that the Mapper is an optimal estimator of the Reeb graph, which gives, as a byproduct, a method to automatically tune its parameters and compute confidence regions on its topological features, such as its loops and flares. This allows to circumvent the issue of testing a large grid of parameters and keeping the most stable ones in the brute-force setting, which is widely used in visualization, clustering and feature selection with the Mapper.

7.2.5. Sliced Wasserstein Kernel for Persistence Diagrams

Participants: Steve Oudot, Mathieu Carrière.

In collaboration with M. Cuturi (ENSAE)

Persistence diagrams (PDs) play a key role in topological data analysis (TDA), in which they are routinely used to describe succinctly complex topological properties of complicated shapes. PDs enjoy strong stability properties and have proven their utility in various learning contexts. They do not, however, live in a space naturally endowed with a Hilbert structure and are usually compared with specific distances, such as the bottleneck distance. To incorporate PDs in a learning pipeline, several kernels have been proposed for PDs with a strong emphasis on the stability of the RKHS distance w.r.t. perturbations of the PDs. In [27], we use the Sliced Wasserstein approximation of the Wasserstein distance to define a new kernel for PDs, which is not only provably stable but also provably discriminative w.r.t. the Wasserstein distance $W1\infty$ between PDs. We also demonstrate its practicality, by developing an approximation technique to reduce kernel computation time, and show that our proposal compares favorably to existing kernels for PDs on several benchmarks.

7.2.6. An introduction to Topological Data Analysis: fundamental and practical aspects for data scientists

Participants: Frédéric Chazal, Bertrand Michel.

Topological Data Analysis (TDA) is a recent and fast growing field providing a set of new topological and geometric tools to infer relevant features for possibly complex data. In [45], we propose a brief introduction, through a few selected recent and state-of-the-art topics, to basic fundamental and practical aspects of TDA for non experts.

7.3. Topological approach for multimodal data processing

7.3.1. On the Stability of Functional Maps and Shape Difference Operators

Participants: Frédéric Chazal, Ruqi Huang, Maks Ovsjanikov.

In this paper, we provide stability guarantees for two frameworks that are based on the notion of functional maps. We consider two types of perturbations in our analysis: one is on the input shapes and the other is on the change in *scale*. In theory, we formulate and justify the robustness that has been observed in practical implementations of those frameworks. Inspired by our theoretical results, we propose a pipeline for constructing shape difference operators on point clouds and show numerically that the results are robust and informative. In particular, we show that both the shape difference operators and the derived areas of highest distortion are stable with respect to changes in shape representation and change of scale. Remarkably, this is in contrast with the well-known instability of the eigenfunctions of the Laplace-Beltrami operator computed on point clouds compared to those obtained on triangle meshes.

7.3.2. Local Equivalence and Intrinsic Metrics Between Reeb Graphs

Participants: Steve Oudot, Mathieu Carrière.

As graphical summaries for topological spaces and maps, Reeb graphs are common objects in the computer graphics or topological data analysis literature. Defining good metrics between these objects has become an important question for applications, where it matters to quantify the extent by which two given Reeb graphs differ. Recent contributions emphasize this aspect, proposing novel distances such as functional distortion or interleaving that are provably more discriminative than the so-called bottleneck distance, being true metrics whereas the latter is only a pseudo-metric. Their main drawback compared to the bottleneck distance is to be comparatively hard (if at all possible) to evaluate. In [28] we take the opposite view on the problem and show that the bottleneck distance is in fact good enough locally, in the sense that it is able to discriminate a Reeb graph from any other Reeb graph in a small enough neighborhood, as efficiently as the other metrics do. This suggests considering the intrinsic metrics induced by these distances, which turn out to be all globally equivalent. This novel viewpoint on the study of Reeb graphs has a potential impact on applications, where one may not only be interested in discriminating between data but also in interpolating between them.

7.3.3. *Structure and Stability of the One-Dimensional Mapper*

Participants: Steve Oudot, Mathieu Carrière.

Given a continuous function $f : X -> R$ and a cover I of its image by intervals, the Mapper is the nerve of a refinement of the pullback cover $f^{-1}(I)$. Despite its success in applications, little is known about the structure and stability of this construction from a theoretical point of view. As a pixelized version of the Reeb graph of f, it is expected to capture a subset of its features (branches, holes), depending on how the interval cover is positioned with respect to the critical values of the function. Its stability should also depend on this positioning. In [16] we propose a theoretical framework relating the structure of the Mapper to that of the Reeb graph, making it possible to predict which features will be present and which will be absent in the Mapper given the function and the cover, and for each feature, to quantify its degree of (in-)stability. Using this framework, we can derive guarantees on the structure of the Mapper, on its stability, and on its convergence to the Reeb graph as the granularity of the cover I goes to zero.

7.4. Experimental research and software development

7.4.1. *Stride detection for pedestrian trajectory reconstruction: a machine learning approach based on geometric patterns*

Participants: Frédéric Chazal, Bertrand Michel, Bertrand Beaufils.

In collaboration with M. Grelet (Sysnav)

A strides detection algorithm is proposed using inertial sensors worn on the ankle. This innovative approach based on geometric patterns can detect both normal walking strides and atypical strides such as small steps, side steps and backward walking that existing methods struggle to detect. It is also robust in critical situations, when for example the wearer is sitting and moving the ankle, while most algorithms in the literature would wrongly detect strides.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- Collaboration with Sysnav, a French SME with world leading expertise in navigation and geopositioning in extreme environments, on TDA, geometric approaches and machine learning for the analysis of movements of pedestrians and patients equipped with inetial sensors (CIFRE PhD of Bertrand Beaufils).
- Collaboration with Fujitsu on TDA and Machine learning (started in Dec 2017).

8.2. Bilateral Grants with Industry

- DATASHAPE and Sysnav have been selected for the ANR/DGA Challenge MALIN (funding: 700 kEuros) in September 2017.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. *ANR*

9.1.1.1. ANR TOPDATA

Participants: Jean-Daniel Boissonnat, Frédéric Chazal, David Cohen-Steiner, Mariette Yvinec, Steve Oudot, Marc Glisse.

- Acronym : TopData.

- Type : ANR blanc.

- Title : Topological Data Analysis: Statistical Methods and Inference.

- Coordinator : Frédéric Chazal (DATASHAPE).

- Duration : 4 years from October 2013 to September 2017.

- Others Partners: Département de Mathématiques (Université Paris Sud), Institut de Mathématiques (Université de Bourgogne), LPMA (Université Paris Diderot), LSTA (Université Pierre et Marie Curie).

- Abstract: TopData aims at designing new mathematical frameworks, models and algorithmic tools to infer and analyze the topological and geometric structure of data in different statistical settings. Its goal is to set up the mathematical and algorithmic foundations of Statistical Topological and Geometric Data Analysis and to provide robust and efficient tools to explore, infer and exploit the underlying geometric structure of various data.

Our conviction, at the root of this project, is that there is a real need to combine statistical and topological/geometric approaches in a common framework, in order to face the challenges raised by the inference and the study of topological and geometric properties of the wide variety of larger and larger available data. We are also convinced that these challenges need to be addressed both from the mathematical side and the algorithmic and application sides. Our project brings together in a unique way experts in Statistics, Geometric Inference and Computational Topology and Geometry. Our common objective is to design new theoretical frameworks and algorithmic tools and thus to contribute to the emergence of a new field at the crossroads of these domains. Beyond the purely scientific aspects we hope this project will help to give birth to an active interdisciplinary community. With these goals in mind we intend to promote, disseminate and make our tools available and useful for a broad audience, including people from other fields.

- See also: http://geometrica.saclay.inria.fr/collaborations/TopData/Home.html

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

9.2.1.1. GUDHI

Title: Algorithmic Foundations of Geometry Understanding in Higher Dimensions

Programm: FP7

Type: ERC

Duration: February 2014 - January 2019

Coordinator: Inria

Inria contact: Jean-Daniel Boissonnat.

'The central goal of this proposal is to settle the algorithmic foundations of geometry understanding in dimensions higher than 3. We coin the term geometry understanding to encompass a collection of tasks including the computer representation and the approximation of geometric structures, and the inference of geometric or topological properties of sampled shapes. The need to understand geometric structures is ubiquitous in science and has become an essential part of scientific computing and data analysis. Geometry understanding is by no means limited to three dimensions. Many applications in physics, biology, and engineering require a keen understanding of the geometry of a variety of higher dimensional spaces to capture concise information from the underlying often highly nonlinear structure of data. Our approach is complementary to manifold learning techniques and aims at developing an effective theory for geometric and topological data analysis. To reach these objectives, the guiding principle will be to foster a symbiotic relationship between theory and practice, and to address fundamental research issues along three parallel advancing fronts. We will simultaneously develop mathematical approaches providing theoretical guarantees, effective algorithms that are amenable to theoretical analysis and rigorous experimental validation, and perennial software development. We will undertake the development of a high-quality open source software platform to implement the most important geometric data structures and algorithms at the heart of geometry understanding in higher dimensions. The platform will be a unique vehicle towards researchers from other fields and will serve as a basis for groundbreaking advances in scientific computing and data analysis.'

9.3. International Initiatives

9.3.1. Inria Associate Teams Not Involved in an Inria International Labs

9.3.1.1. CATS

Title: Computations And Topological Statistics

International Partner (Institution - Laboratory - Researcher):

Carnegie Mellon University (United States) - Department of Statistics - Larry Wasserman

Start year: 2015

See also: http://geometrica.saclay.inria.fr/collaborations/CATS/CATS.html

Topological Data Analysis (TDA) is an emergent field attracting interest from various communities, that has recently known academic and industrial successes. Its aim is to identify and infer geometric and topological features of data to develop new methods and tools for data exploration and data analysis. TDA results mostly rely on deterministic assumptions which are not satisfactory from a statistical viewpoint and which lead to a heuristic use of TDA tools in practice. Bringing together the strong expertise of two groups in Statistics (L. Wasserman's group at CMU) and Computational Topology and Geometry (Inria Geometrica), the main objective of CATS is to set-up the mathematical foundations of Statistical TDA, to design new TDA methods and to develop efficient and easy-to-use software tools for TDA.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

Ramsay Dyer, Mathematical Sciences Publishers, Canada (June and November 2017)

Arijit Ghosh, Indian Statistical Institute, Kolkata (June and november 2017)

Kim Jisu, CMU, Pittsburgh, USA (November 2017).

Wolfgang Polonik, UC Davis, USA (June 2017).

Konstantin Mischaikow, Rutgers University, USA, (November 2017).

Magnus Botnan, TU Munich, Germany (March 2017).

Sara Kalisnik, MPI, Germany (November 2017).

9.4.1.1. Internships

Divyansh Pareek, IIT Bombay (May-July 2017)

9.4.2. Visits to International Teams

9.4.2.1. Research Stays Abroad

Vincent Divol, UC Davis (April-June 2017)

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. Member of the Organizing Committees

- Frédéric Chazal co-organized, with M. Meila 5Univ. of Washington) the NIPS 2017 workshop "Synergies in Geometric Data Analysis", December 2017.
- Frédéric Chazal co-organized the workshop "Functoriality in Geometric Data", Schloss Dagstuhl, Germany, January 2017

10.1.2. Journal

10.1.2.1. Member of the Editorial Boards

Jean-Daniel Boissonnat is a member of the Editorial Board of *Journal of the ACM*, *Discrete and Computational Geometry*, *International Journal on Computational Geometry and Applications*.

Frédéric Chazal is a member of the Editorial Board of *SIAM Journal on Imaging Sciences*, *Discrete and Computational Geometry (Springer)*, *Graphical Models (Elsevier), and Journal of Applied and Computational Topology (Springer)*.

Steve Oudot is a member of the Editorial Board of *Journal of Computational Geometry*.

10.1.3. Invited Talks

Jean-Daniel Boissonnat has been invited and gave a talk at NYU-AD on November 19, 2017.

Frédéric Chazal, Foundations of Computational Mathematics (FoCM'17), Computational Topology and Geometry workshop, Barcelona, Spain, July 2017.

Frédéric Chazal, Applied Topology in Bedlewo 2017 Conference, Bedlewo, Poland, June 2017.

Frédéric Chazal, UC Davis Statistical Sciences Symposium, Davis, USA, May 2017.

Frédéric Chazal, Applied and Computational Algebraic Topology, Hausdorff Institute, Bonn,Germany, April 2017.

Frédéric Chazal, The First International Conference on the Mathematics of Data Science, Hong Kong Baptist University, Hong Kong, March 2017.

Frédéric Chazal, CNA/Ki-Net Workshop: Dynamics and Geometry from High Dimensional Data, Carnegie Mellon University, March 2017.

Frédéric Chazal, Colloquium, collaborative research center Discretization in Geometry and Dynamics, Munich, February 7, 2017.

Frédéric Chazal, Statistics/Learning at Paris-Saclay, workshop at IHES, January 19, 2017.

Mathijs Wintraecken gave an invited talk at the SoCG workshop on Algorithms for the Medial Axis in Brisbane, June 2017.

Steve Oudot, BIRS workshop on Topological Data Analysis: developping abstract foundations, Banff, Canada, August 2017.

Steve Oudot, Mini-Symposium on Computational Topology, Brisbane, Australia, July 2017.

Steve Oudot, Dagstuhl seminar on Topology, computation and data analysis, Dagstuhl, Germany, July 2017.

Steve Oudot, Applied Topology Seminar, Brown University, USA, April 2017.

Steve Oudot, TRIPODS wokshop on Geometry and topology for data, ICERM, USA, December 2017.

Steve Oudot, workshop on Mathematical signal processing and data analysis, Bremen University, Germany, September 2017.

Steve Oudot, Conférence de rentrée Maths-Info, ENS Cachan, September 2017.

10.1.4. Leadership within the Scientific Community

Steve Oudot is co-organizing the monthly seminar on combinatorial and computational geometry at Institut Henri Poincaré.

Steve Oudot is co-head of the GT Géométrie Algorithmique since September 2017.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Collège de France : Jean-Daniel Boissonnat, Géométrie algorithmique: données, modèles, programmes, avril-juin 2017.

Master: Frédéric Chazal, Analyse Topologique des Données, 30h eq-TD, Université Paris-Sud, France.

Master: Jean-Daniel Boissonnat and Marc Glisse, Computational Geometry Learning, 36h eq-TD, M2, MPRI, France.

Master: Steve Oudot, Topological Data Analysis, 45h eq-TD, M1, École Polytechnique, France.

Master: Frédéric Cazals and Frédéric Chazal, Geometric Methods for Data Analysis, 30h eq-TD, M1, École Centrale Paris, France.

Master: Frédéric Chazal, Topological Data Analysis, 16h eq-TD, M2, Mathématiques, Vision, Apprentissage (MVA), ENS Paris-Saclay, France.

Winter School on Computational geometry and topology for data analysis, Jean-Daniel Boissonnat, Frédéric Chazal, Sophia-Antipolis, january 2017.

10.2.2. Supervision

PhD: Eddie Aamari, A Statistical Approach of Topological Data Analysis, September 1st, 2017, Frédéric Chazal (co-advised by Pascal Massart).

PhD in progress: Claire Brécheteau, Statistical aspects of distance-like functions , started September 1st, 2015, Frédéric Chazal (co-advised by Pascal Massart).

PhD in progress: Bertrand Beaufils, Méthodes topologiques et apprentissage statistique pour l'actimétrie du piéton à partir de données de mouvement, started November 2016, Frédéric Chazal (co-advised by Bertrand Michel).

PhD: Mathieu Carrière, Topological signatures for geometric data, defended November 21st, 2017, Steve Oudot.

PhD in progress: Jérémy Cochoy, Decomposition and stability of multidimensional persistence modules, started September 1st, 2015, Steve Oudot.

PhD in progress: Nicolas Berkouk, Categorification of topological graph structures, started November 1st, 2016, Steve Oudot.

PhD in progress: Théo Lacombe, Statistics for persistence diagrams using optimal transport, started October 1st, 2017, Steve Oudot.

PhD in progress: Alba Chiara de Vitis, Concentration of measure and clustering, Jean-Daniel Boissonnat and David Cohen-Steiner.

PhD in progress: Siargey Kachanovich, Manifold reconstruction in higher dimensions, Jean-Daniel Boissonnat.

PhD in progress: François Godi, Data structures and algorithms for topological data analysis and high dimensional geometry, Jean-Daniel Boissonnat.

PhD in progress: Siddharth Pritam, Approximation algorithms in Computational Topology, Jean-Daniel Boissonnat.

PhD in progress: Raphaël Tinarrage, Persistence and stability of nerves in measured metric spaces for Topological Data Analysis, started September 1st, 2017, Frédéric Chazal and Marc Glisse.

PhD in progress: Vincent Divol, statistical aspects of TDA, started September 1st, 2017, Frédéric Chazal (co-advised by Pascal Massart).

10.2.3. Juries

Frédéric Chazal was a member of the PhD defense committee of Aruni Choudhary, MPI (reviewer) and Aurelien Vasseur (co-advisor).

Jean-Daniel Boissonnat was a member of the PhD defense committee of Eddie Amari.

Steve Oudot was a member (examiner) of the PhD defence committees of Jérémy Dubut and Nicolas Ninin (Cosynus team, Ecole polytechnique).

10.3. Popularization

10.3.1. Inria-Industry Meeting

Marc Glisse, Miro Kramar and Steve Oudot held a booth for half a day.

Marc Glisse played for a small video which is now on the InriaInnovation YouTube channel https://youtu.be/lKNjGk-Z6b4.

11. Bibliography

Major publications by the team in recent years

[1] D. ATTALI, U. BAUER, O. DEVILLERS, M. GLISSE, A. LIEUTIER.*Homological Reconstruction and Simplification in R3*, in "Computational Geometry", 2014 [*DOI :* 10.1016/J.COMGEO.2014.08.010], https://hal.archives-ouvertes.fr/hal-01132440.

[2] J. BOISSONNAT, R. DYER, A. GHOSH.*Stability of Delaunay-type structures for manifolds: [extended abstract]*, in "Symposium on Computational Geometry 2012, SoCG '12, Chapel Hill, NC, USA, June 17-20, 2012", 2012, p. 229–238, http://doi.acm.org/10.1145/2261250.2261284.

[3] J.-D. BOISSONNAT, A. GHOSH.*Manifold reconstruction using tangential Delaunay complexes*, in "Discrete and Computational Geometry", January 2014, vol. 51, n^{o} 1, p. 221-267 [*DOI :* 10.1007/s00454-013-9557-2], https://hal.inria.fr/hal-00932209.

[4] J. BOISSONNAT, C. MARIA.*The Simplex Tree: An Efficient Data Structure for General Simplicial Complexes*, in "Algorithmica", 2014, vol. 70, n^{o} 3, p. 406–427, http://dx.doi.org/10.1007/s00453-014-9887-3.

[5] F. CHAZAL, D. COHEN-STEINER, M. GLISSE, L. J. GUIBAS, S. OUDOT.*Proximity of Persistence Modules and Their Diagrams*, in "Proc. 25th Annual Symposium on Computational Geometry", 2009, p. 237–246 [*DOI :* 10.1145/1542362.1542407], http://hal.inria.fr/inria-00292566/.

[6] F. CHAZAL, D. COHEN-STEINER, A. LIEUTIER.*A Sampling Theory for Compact Sets in Euclidean Space*, in "Discrete Comput. Geom.", 2009, vol. 41, n^{o} 3, p. 461–479, http://dx.doi.org/10.1007/s00454-009-9144-8.

[7] F. CHAZAL, D. COHEN-STEINER, Q. MÉRIGOT.*Geometric Inference for Measures based on Distance Functions*, in "Foundations of Computational Mathematics", 2011, vol. 11, n^{o} 6, p. 733-751, RR-6930 [*DOI :* 10.1007/s10208-011-9098-0], http://hal.inria.fr/inria-00383685.

[8] F. DE GOES, D. COHEN-STEINER, P. ALLIEZ, M. DESBRUN.*An Optimal Transport Approach to Robust Reconstruction and Simplification of 2D Shapes*, in "Computer Graphics Forum", 2011, vol. 30, n^{o} 5, p. 1593–1602, Special issue for EUROGRAPHICS Symposium on Geometry Processing, http://hal.inria.fr/hal-00758019.

[9] L. J. GUIBAS, S. Y. OUDOT, P. SKRABA, F. CHAZAL.*Persistence-Based Clustering in Riemannian Manifolds*, in "Journal of the ACM", November 2013, vol. 60, n$^{\text{o}}$ 6, 38, http://hal.archives-ouvertes.fr/hal-00923563.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[10] E. AAMARI.*Convergence Rates for Geometric Inference*, Université Paris-Saclay, September 2017, https://hal.inria.fr/tel-01607782.

[11] M. CARRIERE.*On metric and statistical properties of topological descriptors for geometric data*, Université Paris 11, November 2017, https://hal.archives-ouvertes.fr/tel-01659347.

Articles in International Peer-Reviewed Journal

[12] J.-D. BOISSONNAT, R. DYER, A. GHOSH.*Delaunay Triangulation of Manifolds*, in "Foundations of Computational Mathematics", 2017, vol. 45, 38 [*DOI :* 10.1007/S10208-017-9344-1], https://hal.inria.fr/hal-01509888.

[13] J.-D. BOISSONNAT, R. DYER, A. GHOSH, M. NIKOLAY.*An Obstruction to Delaunay Triangulations in Riemannian Manifolds*, in "Discrete and Computational Geometry", 2017 [*DOI :* 10.1145/336154.336221], https://hal.inria.fr/hal-01583073.

[14] J.-D. BOISSONNAT, R. DYER, A. GHOSH, S. Y. OUDOT.*Only distances are required to reconstruct submanifolds*, in "Computational Geometry", 2017, vol. 66, p. 32 - 67 [*DOI :* 10.1016/J.COMGEO.2017.08.001], https://hal.inria.fr/hal-01583086.

[15] D. BREMNER, O. DEVILLERS, M. GLISSE, S. LAZARD, G. LIOTTA, T. MCHEDLIDZE, G. MOROZ, S. WHITESIDES, S. WISMATH.*Monotone Simultaneous Paths Embeddings in* $\mathbb{R}^d$, in "Discrete Mathematics and Theoretical Computer Science", January 2018, vol. 20, n$^{\text{o}}$ 1, p. 1-11, https://hal.inria.fr/hal-01529154.

[16] M. CARRIERE, S. Y. OUDOT.*Structure and Stability of the One-Dimensional Mapper*, in "Foundations of Computational Mathematics", October 2017, p. 1–64 [*DOI :* 10.1007/S10208-017-9370-Z], https://hal.archives-ouvertes.fr/hal-01633101.

[17] F. CHAZAL, B. FASY, F. LECCI, B. MICHEL, A. RINALDO, L. WASSERMAN.*Robust Topological Inference: Distance To a Measure and Kernel Distance*, in "Journal of Machine Learning Research", 2017, https://arxiv.org/abs/1412.7197v1 , https://hal.inria.fr/hal-01232217.

[18] F. CHAZAL, R. HUANG, M. OVSJANIKOV.*On the Stability of Functional Maps and Shape Difference Operators*, in "Computer Graphics Forum", 2017, p. 1 - 12 [*DOI :* 10.1111/CGF.13238], https://hal.inria.fr/hal-01668186.

[19] M. MANDAD, D. COHEN-STEINER, L. KOBBELT, P. ALLIEZ, M. DESBRUN.*Variance-Minimizing Transport Plans for Inter-surface Mapping*, in "ACM Transactions on Graphics", 2017, vol. 36, 14 [*DOI :* 10.1145/3072959.3073671], https://hal.inria.fr/hal-01519006.

[20] N. MUSTAFA, K. DUTTA, A. GHOSH.*A Simple Proof of Optimal Epsilon Nets*, in "Combinatorica", 2017, https://hal.archives-ouvertes.fr/hal-01360452.

International Conferences with Proceedings

[21] J.-D. Boissonnat, K. Dutta, A. Ghosh, S. Kolay.*Kernelization of the Subset General Position problem in Geometry*, in "MFCS 2017 - 42nd International Symposium on Mathematical Foundations of Computer Science", Alborg, Denmark, August 2017 [*DOI :* 10.4230/LIPIcs.MFCS.2017.25], https://hal.inria.fr/hal-01583101.

[22] J.-D. Boissonnat, K. Dutta, A. Ghosh, S. Kolay.*Tight Kernels for Covering and Hitting: Point Hyperplane Cover and Polynomial Point Hitting Set*, in "LATIN 2018 - 13th Latin American Theoretical INformatics Symposium", Buenos Aires, Argentina, April 2018, https://hal.inria.fr/hal-01669884.

[23] J.-D. Boissonnat, M. Rouxel-Labbé, M. Wintraecken.*Anisotropic triangulations via discrete Riemannian Voronoi diagrams*, in "Symposium on Computational Geometry SoCG 2017", Brisbane, Australia, July 2017 [*DOI :* 10.4230/LIPIcs.SoCG.2017.19], https://hal.inria.fr/hal-01507111.

[24] J.-D. Boissonnat, K. C. Srikanta.*An Efficient Representation for Filtrations of Simplicial Complexes*, in "Symposium on Discrete Algorithms SODA 2017", Barcelona, France, January 2017, https://hal.inria.fr/hal-01416683.

[25] F. Chazal, B. Beaufils, M. Grelet, B. Michel.*Stride detection for pedestrian trajectory reconstruction: a machine learning approach based on geometric patterns*, in "IPIN 2017 - 8th International Conference on Indoor Positioning and Indoor Navigation", Sapporo, Japan, IEEE, September 2017, p. 1-6 [*DOI :* 10.1109/IPIN.2017.8115867], https://hal.inria.fr/hal-01664659.

[26] F. Chazal, M. Ovsjanikov, E. Corman, M. Bronstein, E. Rodola, M. Ben-Chen, L. Guibas, A. Bronstein.*Computing and Processing Correspondences with Functional Maps* , in "ACM SIGGRAPH 2017 Courses", Los Angeles, United States, July 2017, p. 1-62, https://hal.inria.fr/hal-01664767.

Conferences without Proceedings

[27] M. Carriere, M. Cuturi, S. Y. Oudot.*Sliced Wasserstein Kernel for Persistence Diagrams*, in "ICML 2017 - Thirty-fourth International Conference on Machine Learning", Sydney, Australia, August 2017, p. 1-10, https://hal.archives-ouvertes.fr/hal-01633105.

[28] M. Carriere, S. Y. Oudot.*Local Equivalence and Intrinsic Metrics Between Reeb Graphs*, in "SoCG 2017 - 33rd International Symposium on Computational Geometry", Brisbane, Australia, July 2017, p. 1-15 [*DOI :* 10.4230/LIPIcs.SoCG.2017.25], https://hal.archives-ouvertes.fr/hal-01633109.

[29] K. Dutta, A. Ghosh, B. Jartoux, N. Mustafa.*Shallow packings, semialgebraic set systems, Macbeath regions and polynomial partitioning*, in "33rd International Symposium on Computational Geometry (SoCG 2017)", Brisbane, Australia, July 2017, https://hal.archives-ouvertes.fr/hal-01360443.

Scientific Books (or Scientific Book chapters)

[30] J.-D. Boissonnat, F. Chazal, M. Yvinec.*Geometric and Topological Inference*, Cambridge University Press, October 2017, To appear in the red series Cambridge Texts in Applied Mathematics, https://hal.inria.fr/hal-01615863.

[31] F. Chazal, D. Cohen-Steiner, A. Lieutier, Q. Mérigot, B. Thibert.*Inference of curvature using tubular neighborhoods*, in "Lecture Notes in Mathematics", L. Najman, P. Romon (editors), Modern

Approaches to Discrete Curvature, Springer, 2017, vol. 2184, p. 133-158 [*DOI :* 10.1007/978-3-319-58002-9_4], https://hal.archives-ouvertes.fr/hal-01425558.

Research Reports

[32] J.-D. BOISSONNAT, M. ROUXEL-LABBÉ, M. WINTRAECKEN.*Anisotropic triangulations via discrete Riemannian Voronoi diagrams*, Inria Sophia Antipolis, 2017, n° RR-9056, https://hal.inria.fr/hal-01507273.

[33] O. DEVILLERS, M. GLISSE.*Delaunay triangulation of a random sample of a good sample has linear size*, Inria Saclay Ile de France ; Inria Nancy - Grand Est, July 2017, n° RR-9082, 6, https://hal.inria.fr/hal-01568030.

[34] M. ROUXEL-LABBÉ, M. WINTRAECKEN, J.-D. BOISSONNAT.*Discretized Riemannian Delaunay Triangulations*, Inria Sophia Antipolis - Méditerranée, October 2017, n° RR-9103, 51, https://hal.inria.fr/hal-01612924.

Other Publications

[35] E. AAMARI, J. KIM, F. CHAZAL, B. MICHEL, A. RINALDO, L. WASSERMAN.*Estimating the Reach of a Manifold*, May 2017, https://arxiv.org/abs/1705.04565 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01521955.

[36] E. AAMARI, C. LEVRARD.*Non-Asymptotic Rates for Manifold, Tangent Space, and Curvature Estimation*, April 2017, https://arxiv.org/abs/1705.00989 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01516032.

[37] E. AAMARI, C. LEVRARD.*Stability and Minimax Optimality of Tangential Delaunay Complexes for Manifold Reconstruction*, November 2017, https://arxiv.org/abs/1512.02857 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01245479.

[38] J.-D. BOISSONNAT, O. DEVILLERS, K. DUTTA, M. GLISSE.*Delaunay triangulation of a random sample of a good sample has linear size* , December 2017, working paper or preprint, https://hal.inria.fr/hal-01673170.

[39] J.-D. BOISSONNAT, K. DUTTA, A. GHOSH, S. KOLAY.*Tight Kernels for Covering with Points and Polynomials*, May 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01518562.

[40] J.-D. BOISSONNAT, R. DYER, A. GHOSH, M. WINTRAECKEN.*Local criteria for triangulation of manifolds* *, December 2017, A full version of the paper is available at [5], http://ramsaydyer.com/tmp/loccrit.pdf [*DOI :* 10.4230/LIPICS.CVIT.2016.23], https://hal.inria.fr/hal-01661230.

[41] J.-D. BOISSONNAT, A. LIEUTIER, M. WINTRAECKEN.*The reach, metric distortion, geodesic convexity and the variation of tangent spaces*, December 2017, working paper or preprint, https://hal.inria.fr/hal-01661227.

[42] J.-D. BOISSONNAT, M. WINTRAECKEN.*Triangulating stratified manifolds I: a reach comparison theorem*, December 2017, working paper or preprint, https://hal.inria.fr/hal-01661233.

[43] C. BRÉCHETEAU.*The DTM-signature for a geometric comparison of metric-measure spaces from samples*, January 2017, working paper or preprint, https://hal.inria.fr/hal-01426331.

[44] M. CARRIERE, B. MICHEL, S. Y. OUDOT.*Statistical analysis and parameter selection for Mapper*, November 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01633106.

[45] F. CHAZAL, B. MICHEL.*An introduction to Topological Data Analysis: fundamental and practical aspects for data scientists*, October 2017, https://arxiv.org/abs/1710.04019 - working paper or preprint, https://hal.inria.fr/hal-01614384.

[46] A. CHOUDHARY, S. KACHANOVICH, M. WINTRAECKEN.*Coxeter triangulations have good quality*, December 2017, working paper or preprint, https://hal.inria.fr/hal-01667404.

[47] D. COHEN-STEINER, W. KONG, C. SOHLER, G. VALIANT.*Approximating the Spectrum of a Graph*, December 2017, https://arxiv.org/abs/1712.01725 - working paper or preprint, https://hal.inria.fr/hal-01661199.

[48] K. DUTTA, A. GHOSH.*On Subgraphs of Bounded Degeneracy in Hypergraphs*, December 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01669886.

Project-Team DEDUCTEAM

DEDUCTEAM

IN PARTNERSHIP WITH:
ENS Paris-Saclay

RESEARCH CENTER
Saclay - Île-de-France

THEME
Proofs and Verification

Table of contents

Project-Team DEDUCTEAM

Creation of the Team: 2011 December 01, updated into Project-Team: 2017 January 01

Keywords:

Computer Science and Digital Science:

A2.1.3. - Functional programming
A2.1.11. - Proof languages
A2.4.3. - Proofs
A3.1.1. - Modeling, representation
A7. - Theory of computation
A7.2. - Logic in Computer Science

Other Research Topics and Application Domains:

B7. - Transport and logistics

1. Personnel

Research Scientists

Gilles Dowek [Team leader, Inria, Senior Researcher, HDR]
Frédéric Blanqui [Inria, Researcher, HDR]

Faculty Members

Guillaume Burel [ENSIIE, Associate Professor, delegation Inria since Aug 2017]
Catherine Dubois [ENSIIE, Professor, HDR]
Olivier Hermant [Mines ParisTech, Associate Professor, HDR]

PhD Students

Guillaume Bury [ENS Paris]
Frédéric Gilbert [Ministère de l'Ecologie, de l'Energie, du Développement durable et de la Mer, until Aug 2017]
François Thiré [ENS Paris-Saclay]
Guillaume Genestier [ENS Paris-Saclay since Oct 2017, intern Ecole Polytechnique Apr-Aug 2017]
Gaspard Férey [Mines ParisTech since Sep 2017]

Post-Doctoral Fellows

Rodolphe Lepigre [Inria, from Sep 2017, until Dec 2018]
Simon Martiel [Univ. Paris-Est Créteil, until Mar 2017]

Visiting Scientists

Jean-Pierre Jouannaud [Prof. Emeritus, Univ. Paris-Saclay, Ecole Polytechnique, HDR]
Chaitanya Leena Subramaniam [Intern Ecole Polytechnique, Mar-Aug 2017]
Rafaël Bocquet [Intern ENS Paris Mar-Aug 2017]
Antoine Defourné [Intern ENSIMAG Feb-Aug 2017]

Administrative Assistants

Thida Iem [Inria until Jul 2017]
Emmanuelle Perrot [Inria since Jul 2017]

2. Overall Objectives

2.1. Objectives

The project-team investigates applications of proof theory to the design of logical frameworks, to interoperability between proof systems, and to the development of system-independent proof libraries.

To achieve these goals, we develop a logical framework DEDUKTI, where several theories can be expressed, systems translating proofs between external proof systems and DEDUKTI theories, tools to migrate proofs within DEDUKTI from one theory to another, tools to prove the confluence, the termination, and the consistency of theories expressed in DEDUKTI, and tools to develop proofs directly in DEDUKTI.

2.2. History

Deduction modulo theory is a formulation of predicate logic where deduction is performed modulo an equivalence relation defined on propositions. A typical example is the equivalence relation relating propositions differing only by a re-arrangement of brackets around additions, relating, for instance, the propositions $P((x+y)+z)$ and $P(x+(y+z))$. Reasoning modulo this equivalence relation permits to drop the associativity axiom. Thus, in Deduction modulo theory, a theory is formed with a set of axioms and an equivalence relation. When the set of axioms is empty the theory is called purely computational.

Deduction modulo theory was proposed at the end of the 20th century as a tool to simplify the completeness proof of equational resolution [6]. Soon, it was noticed that this idea was also present in other areas of logic, such as Martin-Löf's type theory, where the equivalence relation is definitional equality, Prawitz' extended natural deduction, etc. [2]. More generally, Deduction modulo theory gives an account on the way reasoning and computation are articulated in a formal proof, a topic slightly neglected by logic, but of prime importance when proofs are computerized.

The early research on Deduction modulo theory focused on the design of general proof search methods—Resolution modulo theory, tableaux modulo theory, etc.—that could be applied to any theory formulated in Deduction modulo theory, to general proof normalization and cut elimination results, to the definitions of models taking the difference between reasoning and computation into account, and to the definition of specific theories—simple type theory, arithmetic, some versions of set theory, etc.—as purely computational theories.

A new turn with Deduction modulo theory was taken when the idea of reasoning modulo an arbitrary equivalence relation was applied to typed λ-calculi with dependent types, that permits to express proofs as algorithms, using the Brouwer-Heyting-Kolmogorov interpretation and the Curry-de Bruijn-Howard correspondence [5]. It was shown in 2007, that extending the simplest λ-calculus with dependent types, the $\lambda\Pi$-calculus, with an equivalence relation (more precisely a congruence), led to a calculus, called the $\lambda\Pi$-calculus modulo theory, that permits to simulate many other λ-calculi, such as the Calculus of Constructions, designed to express proofs in specific theories.

This led to the development of a logical framework based on the $\lambda\Pi$-calculus modulo theory [29], that could be used to verify proofs coming from different proof systems, such as COQ [27], HOL [32], etc. To emphasize the versatility of this system, we called it DEDUKTI—"to deduce" in Esperanto. This system is currently developed together with companion systems, COQINE, KRAJONO, HOLIDE, FOCALIDE, and ZENONIDE, that permits to translate proofs from COQ, HOL, FOCALIZE, and ZENON, to DEDUKTI. Other tools, such as ZENON MODULO, directly output proofs that can be checked by DEDUKTI. DEDUKTI proofs can also be exported to several other systems. All this is presented in [1].

3. Research Program

3.1. Proof checking

A thesis, which is at the root of our research effort, and which was already formulated in [31], is that proof checkers should be theory independent. This is for instance expressed in the title of our invited talk at ICALP 2012: *A theory independent Curry-De Bruijn-Howard correspondence* [30]. Such a theory independent proof checker is called a logical framework.

Part of our research effort is focused on improving the $\lambda\Pi$-calculus modulo theory, for instance allowing to define congruences with associative and commutative rewriting.

Another part of our research effort is focused on the automatic analysis of theories to prove their confluence, termination, and consistency automatically [3].

3.2. Interoperability

Using a single prover to check proofs coming from different provers naturally leads to investigate how these proofs can interact one with another. This issue is of prime importance because developments in proof systems are getting bigger and, unlike other communities in computer science, the proof checking community has given little effort in the direction of standardization and interoperability.

For each proof, independently of the system in which it has been developed, we should be able to identify the systems in which it can be expressed. For instance, we have shown that many proofs developed in the MATITA prover did not use the full strength of the logic of MATITA and could be exported, for instance, to the systems of the HOL family, that are based on a weaker logic.

3.3. Libraries

Rather than importing proofs from one system, transforming them, and exporting them to another system, we can use the same tools to develop system-independant proof librairies. In such a library, each proof is labeled with the logics in which it can be expressed and so with the systems in which it can be used.

3.4. Interactive theorem proving

If our main goal with DEDUKTI is to import, transform, and export proofs developed in other systems, we want also, in some cases, to develop proofs interactively directly in DEDUKTI. This leads to the development of a tactic system, called DEMON, on top of DEDUKTI.

4. Application Domains

4.1. Safety of aerospace systems

In parallel with this effort in logic and in the development of proof checkers and automated theorem proving systems, we always have been interested in using such tools. One of our favorite application domain is the safety of aerospace systems. Together with César Muñoz' team in Nasa-Langley, we have proved the correctness of several geometric algorithms used in air traffic control.

This has led us sometimes to develop such algorithms ourselves, and sometimes to develop tools for automating these proofs.

4.2. Termination certificate verification

Termination is an important property to verify, especially in critical applications. Automated termination provers use more and more complex theoretical results and external tools (e.g. sophisticated SAT solvers) that make their results not fully trustable and very difficult to check. To overcome this problem, a language for termination certificates, called CPF, has been developed. Deducteam develops a formally certified tool, RAINBOW, based on the Coq library COLOR, that is able to automatically verify the correctness of some of these termination certificates.

5. New Software and Platforms

5.1. Autotheo

KEYWORD: Automated deduction
SCIENTIFIC DESCRIPTION: Transformation of axiomatic theories into rewriting systems that can be used by iProverModulo.
FUNCTIONAL DESCRIPTION: Autotheo is a tool that transforms axiomatic theories into polarized rewriting systems, thus making them usable in iProverModulo. It supports several strategies to orient the axioms, some of them being proved to be complete, in the sense that ordered polarized resolution modulo the resulting systems is refutationally complete, some others being merely heuristics. In practice, Autotheo takes a TPTP input file and produces an input file for iProverModulo.
NEWS OF THE YEAR: Used by iProverModulo in its participation at the CASC-26 competition.

- Participant: Guillaume Burel
- Partner: ENSIIE
- Contact: Guillaume Burel
- Publication: Consistency Implies Cut Admissibility
- URL: http://www.ensiie.fr/~guillaume.burel/blackandwhite_autotheo.html.en

5.2. CoLoR

Coq Library on Rewriting and termination
KEYWORDS: Coq - Formalisation
FUNCTIONAL DESCRIPTION: CoLoR is a Coq library on rewriting theory and termination. It provides many definitions and theorems on various mathematical structures (quasi-ordered sets, relations, ordered semi-rings, etc.), data structures (lists, vectors, matrices, polynomials, finite graphs), term structures (strings, first-order terms, lambda-terms, etc.), transformation techniques (dependency pairs, semantic labeling, etc.) and (non-)termination criteria (polynomial and matrix interpretations, recursive path ordering, computability closure, etc.).
NEWS OF THE YEAR: 2017: Port to Coq 8.6 and 8.7.

- Authors: Frédéric Blanqui and Sébastien Hinderer
- Contact: Frédéric Blanqui
- Publications: CoLoR: a Coq library on well-founded rewrite relations and its application to the automated verification of termination certificates - Automated Verification of Termination Certificates - CoLoR: a Coq library on rewriting and termination
- URL: http://color.inria.fr/

5.3. Coqine

Coq In dEdukti
KEYWORDS: Higher-order logic - Formal methods - Proof
FUNCTIONAL DESCRIPTION: CoqInE is a plugin for the Coq software translating Coq proofs into Dedukti terms. It provides a Dedukti signature file faithfully encoding the underlying theory of Coq (or a sufficiently large subset of it). Current development is mostly focused on implementing support for Coq universe polymorphism. The generated ouput is meant to be type-checkable using the latest version of Dedukti.

- Contact: Guillaume Burel
- URL: http://www.ensiie.fr/~guillaume.burel/blackandwhite_coqInE.html.en

5.4. Dedukti

KEYWORD: Logical Framework
FUNCTIONAL DESCRIPTION: Dedukti is a proof-checker for the LambdaPi-calculus modulo. As it can be parametrized by an arbitrary set of rewrite rules, defining an equivalence relation, this calculus can express many different theories. Dedukti has been created for this purpose: to allow the interoperability of different theories.

Dedukti's core is based on the standard algorithm for type-checking semi-full pure type systems and implements a state-of-the-art reduction machine inspired from Matita's and modified to deal with rewrite rules.

Dedukti's input language features term declarations and definitions (opaque or not) and rewrite rule definitions. A basic module system allows the user to organize his project in different files and compile them separately.

Dedukti features matching modulo beta for a large class of patterns called Miller's patterns, allowing for more rewriting rules to be implemented in Dedukti.

- Participants: François Thiré, Gaspard Ferey, Guillaume Genestier and Rodolphe Lepigre
- Contact: François Thiré
- Publications: Dedukti:un vérificateur de preuves universel - Rewriting Modulo β in the $\lambda\Pi$-Calculus Modulo - Expressing theories in the $\lambda\Pi$-calculus modulo theory and in the Dedukti system
- URL: http://dedukti.gforge.inria.fr/

5.5. Holide

KEYWORD: Proof
FUNCTIONAL DESCRIPTION: Holide translates HOL proofs to Dedukti[OT] proofs, using the OpenTheory standard (common to HOL Light and HOL4). Dedukti[OT] being the encoding of OpenTheory in Dedukti.

- Contact: Guillaume Burel
- URL: http://deducteam.gforge.inria.fr/holide/

5.6. HOT

Higher-Order Termination
FUNCTIONAL DESCRIPTION: HOT is an automated termination prover for higher-order rewriting, based on the notion of computability closure.

- Contact: Frédéric Blanqui
- URL: http://rewriting.gforge.inria.fr/hot.html

5.7. iProver Modulo

KEYWORDS: Automated deduction - Automated theorem proving
SCIENTIFIC DESCRIPTION: Integration of ordered polarized resolution modulo theory into the prover iProver.
FUNCTIONAL DESCRIPTION: iProver Modulo is an extension of the automated theorem prover iProver originally developed by Konstantin Korovin at the University of Manchester. It implements ordered polarized resolution modulo theory, a refinement of the resolution method based on deduction modulo theory. It takes as input a proposition in predicate logic and a clausal rewriting system defining the theory in which the formula has to be proved. Normalization with respect to the term rewriting rules is performed very efficiently through translation into OCaml code, compilation and dynamic linking. Experiments have shown that ordered polarized resolution modulo dramatically improves proof search compared to using raw axioms.

News Of The Year: Participation at the automated-theorem-prover competition CASC-26 Integration of version 2.5 of iProver, adding support for types (TFF0)

- Participant: Guillaume Burel
- Partner: ENSIIE
- Contact: Guillaume Burel
- Publications: A Shallow Embedding of Resolution and Superposition Proofs into the ??-Calculus Modulo - Experimenting with deduction modulo
- URL: http://www.ensiie.fr/~guillaume.burel/blackandwhite_iProverModulo.html.en

5.8. mSAT

Keyword: Propositional logic
Functional Description: mSAT is a modular, proof-producing, SAT and SMT core based on Alt-Ergo Zero, written in OCaml. The solver accepts user-defined terms, formulas and theory, making it a good tool for experimenting. This tool produces resolution proofs as trees in which the leaves are user-defined proof of lemmas.

- Contact: Guillaume Bury
- Publication: mSAT:An OCaml SAT Solver
- URL: https://github.com/Gbury/mSAT

5.9. Rainbow

Termination certificate verifier
Keywords: Demonstration - Code generation - Verification
Functional Description: Rainbow is a set of tools for automatically verifying the correctness of termination certificates expressed in the CPF format used in the annual international competition of termination tools. It contains: a tool xsd2coq for generating Coq data types for representing XML files valid with respect to some XML Schema, a tool xsd2ml for generating OCaml data types and functions for parsing XML files valid with respect to some XML Schema, a tool for translating a CPF file into a Coq script, and a standalone Coq certified tool for verifying the correctness of a CPF file.

- Author: Frédéric Blanqui
- Contact: Frédéric Blanqui
- Publications: Automated verification of termination certificates - Automated verification of termination certificates
- URL: http://color.inria.fr/rainbow.html

5.10. Krajono

Keyword: Proof
Functional Description: Krajono translates Matita proofs into Dedukti[CiC] (encoding of CiC in Dedukti) terms.

- Contact: François Thiré

5.11. archsat

Keywords: Automated theorem proving - First-order logic - Propositional logic
Functional Description: Archsat is an automated theorem prover aimed at studying the integration of first-order theorem prover technologies, such as rewriting, into SMT solvers.

- Contact: Guillaume Bury
- URL: https://gforge.inria.fr/projects/archsat

6. New Results

6.1. $\lambda\Pi$-calculus modulo theory

G. Dowek has given a semantic criterion for the termination of the $\lambda\Pi$-calculus modulo theory. This result has been published in [23].

A. Assaf, G. Dowek, J.-P. Jouannaud and J. Liu have given a confluence criterion for untyped higher-order rewrite systems, and demonstrated some applications to the $\lambda\Pi$-calculus modulo theory.

G. Dowek has given an invited talk at PxTP where he has presented a state of the art of the production of system-independent proof libraries. This paper has been published in the proceedings of PxTP [12].

6.2. Dedukti

During his internship [22], A. Defourné extended F. Blanqui's prototype of proof assistant based on Dedukti by developing a tactic for calling external provers through Why3 [28]. He also started to study a simple rewriting tactic.

During his internship, R. Bocquet studied unification in the $\lambda\Pi$-calculus modulo rewriting, and started to implement a prototype.

During his internship [24], G. Genestier studied the possibility to use the Size-Change Principle [34] in order to prove termination in the $\lambda\Pi$-calculus modulo rewriting. This work led to an adaptation of the criterion developped in his thesis by Wahlstedt [40] to a calculus containing dependant types. He also implemented a prototype of a weak version of the criterion.

During the first three months of his postdoc, R. Lepigre proposed a new implementation of Dedukti [36], based on the Bindlib library for the representation of structures with binders [38]. The libraries generated for Dedukti are compatible with this new implementation, and can be type-checked with minor modifications.

During the first months of his PhD, G. Férey adapted the higher-order pattern matching and convertibility checking algorithms to implemented support for rewriting modulo associative-commutative (AC) symbols in Dedukti.

6.3. Interoperability

F. Thiré has finished to implement a translation of an arithmetic library from Matita to OpenTheory. This work can be decomposed in two steps: A first step goes from Matita to a new logic called STTforall while a second step goes from STTforall to OpenTheory. This translation will be described in two separate papers. The first paper that will be submitted to FSCD 2018 describe the logic STTforall and its translation to HOL while the second paper explains the translation from Matita to STTforall. STTforall is a very simple logic and so, it is easy to translate proofs from this logic to other proofs assistants. For example, a translation from STTforall to Coq has also been implemented by F. Thiré. Two new tools have been implemented to make these translations:

- Dkmeta is a tool that translates terms thanks to the rewrite engine of Dedukti
- Ediloh is a tool that translates terms from STTforall them in OpenTheory

F.Gilbert developed a first prototype for the extraction of proofs from the proof assistant PVS that can be verified externally. The system PVS is based on the dichotomy between a *type-checker* and a *prover*. This proof extraction mechanism is built by instrumenting the PVS *prover*, but does not contain any typing information from the *type-checker* at this stage. Proofs can be built for any PVS theory. However, some reasoning steps rely on unverified assumptions. For a restricted fragment of PVS, the proofs are exported to Dedukti, and the unverified assumptions are proved externally using the automated theorem prover MetiTarski. This work has been published and presented in [15].

6.4. Termination

F. Blanqui revised his paper on "size-based termination of higher-order rewrite systems" submitted to the Journal of Functional Programming [19]. This paper provides a general and modular criterion for the termination of simply-typed λ-calculus extended with function symbols defined by user-defined rewrite rules. Following a work of Hughes, Pareto and Sabry for functions defined with a fixpoint operator and pattern-matching [33], several criteria use typing rules for bounding the height of arguments in function calls. In this paper, we extend this approach to rewriting-based function definitions and more general user-defined notions of size.

R. Lepigre worked on his paper "Practical Subtyping for System F with Sized (Co-)Induction" [39] (joint work with C. Raffalli), which was submitted to the journal Transactions on Programming Languages and Systems (TOPLAS) and is now under revision. This paper proposes a practical type system for a rich, normalizing, extension of (Curry-style) System F. The termination of recursive programs is established using a new mechanism based on circular proofs, which is also used to deal with (sized) inductive and coinductive types (in subtyping). The idea is to build (possibly ill-formed) infinite, circular typing (resp. subtyping) derivations, and to check for their well-foundedness a posteriori. The normalization proof then follows using standard realizability (or reducibility) techniques, the main point being that the adequacy lemma can still be proved by (well-founded) induction on the structure of the "circular" typing (resp. subtyping) derivations.

6.5. Proof theory

G. Burel developed a general framework, focusing with selection, of which various logical systems are instances: ordinary focusing, refinements of resolution, deduction modulo theory, superdeduction and beyond [20]. This strengthens links between sequent calculi and resolution methods.

F. Gilbert developed a constructivization algorithm, taking as input the classical proof of some formula and generating as output, whenever possible, a constructive proof of the same formula. This result has been published and presented in [14].

F. Gilbert submitted his PhD dissertation (work document [25]), centered on the extension of higher-order logic with predicate subtyping. Predicate subtyping is a key feature of the proof assistant PVS, allowing to define types from predicates – for instance, using this feature, the type of even numbers can be defined from the corresponding predicate. The core of this work is the definition of a language of verifiable certificates for predicate subtyping, as well as the proof of two properties of this language: a cut-elimination theorem, a theorem of conservativity over higher-order logic. F. Gilbert presented this language of certificates as well as the cut-elimination theorem at the workshop TYPES 2017.

6.6. Automated theorem proving

G. Bury presented the mSAT library at the OCaml workshop during the International Conference on Functional Programming [21]. This library provides an efficient SAT/SMT solver core written in OCaml, and presented as a functor to allow instantiation with different theories.

6.7. Program verification

R. Lepigre submitted a paper describing the PML_2 programming language and proof assistant [35], which was the main object of his recently defended PhD thesis [37].

6.8. Quantum computing

A. Díaz-Caro and G. Dowek have developed a type system for the λ-calculus that permits to distinguish duplicable terms from non duplicable ones. This work has been presented at Theory and Practice of Natural Computing [13].

7. Partnerships and Cooperations

7.1. National Initiatives

7.1.1. ANR PROGRAMme

This is an ANR for junior researcher Liesbeth Demol (CNRS, UMR 8163 STL, University Lille 3) to which G. Dowek participates. The subject is: "What is a program? Historical and Philosophical perspectives". This project aims at developing the first coherent analysis and pluralistic understanding of "program" and its implications to theory and practice.

7.2. International Initiatives

7.2.1. Participation in Other International Programs

7.2.1.1. International Initiatives

FoQCoSS

Title: Foundations of Quantum Computation: Syntax and Semantics

International Partners (Institution - Laboratory - Researcher):

Universidad Nacional de Quilmes (Argentina) - Alejandro Díaz-Caro

CNRS (France) - Simon Perdrix

Universidade Federal de Santa Maria (Brazil) - Juliana Kaizer Vizzotto

Duration: 2016 - 2017

Start year: 2016

The design of quantum programming languages involves the study of many characteristics of languages which can be seen as special cases of classical systems: parallelism, probabilistic systems, non-deterministic systems, type isomorphisms, etc. This project proposes to study some of these characteristics, which are involved in quantum programming languages, but also have a more immediate utility in the study of nowadays systems. In addition, from a more foundational point of view, we are interested in the implications of computer science principles for quantum physics. For example, the consequences of the Church-Turing thesis for Bell-like experiments: if some of the parties in a Bell-like experiment use a computer to decide which measurements to make, then the computational resources of an eavesdropper have to be limited in order to have a proper observation of non-locality. The final aim is to open a new direction in the search for a framework unifying computer science and quantum physics.

7.3. International Research Visitors

7.3.1. Visits of International Scientists

A. Díaz-Caro (Universidad Nacional de Quilmes, Argentina) visited Deducteam 3 weeks.

7.3.2. Visits to International Teams

7.3.2.1. Research Stays Abroad

F. Thiré has visited the Computation and Logic Group at McGill University for three months.

G. Dowek has visited the university of Quilmes in Buenos Aires for two weeks.

G. Dowek has visited the Pontifical University at Rio for three weeks.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events Selection

8.1.1.1. Member of the Conference Program Committees

G. Burel was a member of the 12th International Workshop on the Implementation of Logics.

8.1.1.2. Reviewer

G. Burel, S. Martiel, and F. Gilbert rewiewed submissions to the Logic In Computer Science conference.

G. Burel reviewed submissions to the Formal Structures for Computation and Deduction conference.

8.1.2. Journal

8.1.2.1. Reviewer - Reviewing Activities

G. Burel reviewed articles for the Computer Journal and the Journal of Logic and Computation.

8.1.3. Invited Talks

G. Burel gave an invited lecture at the 28th Journées Francophones des Langages Applicatifs, entitled “Exprimer ses théories en Dedukti, le vérificateur de preuves universel”.

8.1.4. Scientific Expertise

G. Dowek is a member of the scientific concil of La Société Inforamatique de France.

G. Dowek is a member of the scientific concil of La Main à la Pâte.

G. Dowek is a member of the Commission de réflexion sur l’Éthique de la Recherche en sciences et technologies du Numérique d’Allistene.

8.1.5. Research Administration

F. Blanqui is co-director of the pole 4 (programming: models, algorithms, languages and architectures) of Paris-Saclay University’s doctoral school on computer science.

F. Blanqui is referent of LSV PhD students.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

F. Blanqui gave a M1 course (16h) on rewriting theory in the MPRI at the ENS Paris-Saclay.

F. Blanqui gave a M1 course (21h) on language theory at the ENSIIE.

G. Dowek’s paper *Rules and derivations in an elementary logic course* has been published in the IfCoLog Journal of Logics and their Applications [11].

8.2.2. Supervision

F. Blanqui supervised the internships of A. Defourné and R. Bocquet.

F. Blanqui and O. Hermant supervised the internship of G. Genestier.

F. Blanqui and O. Hermant supervise the PhD of G. Genestier.

G. Dowek supervises the PhD of G. Férey and F. Gilbert.

G. Dowek and D. Delahaye supervise the PhD of G. Bury.

G. Dowek and S. Graham-Lengrand supervise the PhD of F. Thiré.

8.2.3. Juries

F. Blanqui was member of the jury for the best scientific production of the year within Paris-Saclay University's doctoral school on computer science.

F. Blanqui has been in the jury for the PhD of R. Lepigre on “Semantics and Implementation of an Extension of ML for Proving Programs”, Chambéry.

9. Bibliography

Major publications by the team in recent years

[1] A. ASSAF, G. BUREL, R. CAUDERLIER, D. DELAHAYE, G. DOWEK, C. DUBOIS, F. GILBERT, P. HALMAGRAND, O. HERMANT, R. SAILLARD.*Expressing theories in the $\lambda\Pi$-calculus modulo theory and in the Dedukti system*, in "22nd International Conference on Types for Proofs and Programs, TYPES 2016", Novi SAd, Serbia, May 2016, https://hal-mines-paristech.archives-ouvertes.fr/hal-01441751.

[2] F. BLANQUI.*Definitions by rewriting in the Calculus of Constructions*, in "Mathematical Structures in Computer Science", 2005, vol. 15, n^{o} 1, p. 37-92 [*DOI :* 10.1017/S0960129504004426], http://hal.inria.fr/inria-00105648/en/.

[3] F. BLANQUI, J.-P. JOUANNAUD, A. RUBIO.*The Computability Path Ordering*, in "Logical Methods in Computer Science", October 2015 [*DOI :* 10.2168/LMCS-11(4:3)2015], https://hal.inria.fr/hal-01163091.

[4] G. BUREL.*Experimenting with Deduction Modulo*, in "CADE 2011", V. SOFRONIE-STOKKERMANS, N. BJØRNER (editors), Lecture Notes in Artificial Intelligence, Springer, 2011, vol. 6803, p. 162–176.

[5] D. COUSINEAU, G. DOWEK.*Embedding Pure Type Systems in the lambda-Pi-calculus modulo*, in "Typed lambda calculi and applications", S. RONCHI DELLA ROCCA (editor), Lecture Notes in Computer Science, Springer-Verlag, 2007, vol. 4583, p. 102-117.

[6] G. DOWEK, T. HARDIN, C. KIRCHNER.*Theorem proving modulo*, in "Journal of Automated Reasoning", 2003, vol. 31, p. 33-73.

[7] C. DUBOIS, T. HARDIN, V. DONZEAU-GOUGE.*Building certified components within FOCAL*, in "Revised Selected Papers from the Fifth Symposium on Trends in Functional Programming, TFP 2004, München, Germany, 25-26 November 2004", H.-W. LOIDL (editor), Trends in Functional Programming, Intellect, 2006, vol. 5, p. 33-48.

[8] O. HERMANT.*Resolution is Cut-Free*, in "Journal of Automated Reasoning", March 2010, vol. 44, n^{o} 3, p. 245-276.

[9] M. JACQUEL, K. BERKANI, D. DELAHAYE, C. DUBOIS.*Verifying B Proof Rules using Deep Embedding and Automated Theorem Proving*, in "Software and Systems Modeling (SoSyM)", June 2013.

[10] M. JACQUEL, K. BERKANI, D. DELAHAYE, C. DUBOIS.*Tableaux Modulo Theories Using Superdeduction*, in "Global Journal of Advanced Software Engineering (GJASE)", December 2014, vol. 1, p. 1 - 13 [*DOI :* 10.1007/978-3-642-31365-3_26], https://hal.archives-ouvertes.fr/hal-01099338.

Publications of the year

Articles in International Peer-Reviewed Journal

[11] G. DOWEK.*Rules and derivations in an elementary logic course*, in "IfColog Journal of Logics and their Applications (FLAP)", January 2017, vol. 4, n^{o} 1, https://arxiv.org/abs/1601.01483 , https://hal.inria.fr/hal-01252124.

Invited Conferences

[12] G. DOWEK.*Analyzing individual proofs as the basis of interoperability between proof systems*, in "PxTP 2017 - Fifth Workshop on Proof eXchange for Theorem Proving", Brasilia, Brazil, September 2017, https://hal.inria.fr/hal-01670394.

International Conferences with Proceedings

[13] A. DÍAZ-CARO, G. DOWEK.*Typing Quantum Superpositions and Measurement*, in "TPNC 2017 - 6th International Conference on the Theory and Practice of Natural Computing", Prague, Czech Republic, C. MARTÍN-VIDE, R. NERUDA, M. A. VEGA-RODRÍGUEZ (editors), Lecture Notes in Computer Science, Springer, December 2017, vol. 10687, 13, https://arxiv.org/abs/1601.04294 [*DOI :* 10.1007/978-3-319-71069-3_22], https://hal.inria.fr/hal-01670387.

[14] F. GILBERT.*Automated Constructivization of Proofs*, in "FoSSaCS 2017", Uppsala, Sweden, April 2017 [*DOI :* 10.1007/978-3-662-54458-7_28], https://hal.inria.fr/hal-01516788.

[15] F. GILBERT.*Proof certificates in PVS*, in "ITP 2017", Brasilia, Brazil, September 2017 [*DOI :* 10.1007/978-3-319-66107-0_17], https://hal.inria.fr/hal-01673517.

[16] J.-P. JOUANNAUD, P.-Y. STRUB.*Coq without Type Casts: A Complete Proof of Coq Modulo Theory*, in "LPAR-21: 21st International Conference on Logic for Programming, Artificial Intelligence and Reasoning", Maun, Botswana, May 2017, https://hal.inria.fr/hal-01664457.

Scientific Popularization

[17] S. ABITEBOUL, G. DOWEK.*Le temps des algorithmes*, Editions Le Pommier, 2017, 192, https://hal.inria.fr/hal-01502505.

Other Publications

[18] A. ASSAF, G. DOWEK, J.-P. JOUANNAUD, J. LIU.*Untyped Confluence In Dependent Type Theories: Full version*, April 2017, working paper or preprint, https://hal.inria.fr/hal-01515505.

[19] F. BLANQUI.*Size-based termination of higher-order rewriting*, December 2017, working paper or preprint, https://hal.inria.fr/hal-01424921.

[20] G. BUREL.*Linking Focusing and Resolution with Selection*, December 2017, working paper or preprint, https://hal.inria.fr/hal-01670476.

[21] G. BURY.*mSAT:An OCaml SAT Solver*, September 2017, OCaml Users and Developers Workshop, Poster, https://hal.inria.fr/hal-01670765.

[22] A. DEFOURNÉ.*Proof Tactics in Dedukti*, ENSIMAG, September 2017, https://hal.inria.fr/hal-01661872.

[23] G. DOWEK.*Models and termination of proof reduction in the $\lambda\Pi$-calculus modulo theory*, April 2017, working paper or preprint, https://hal.inria.fr/hal-01101834.

[24] G. GENESTIER.*Termination checking in the $\lambda\Pi$-calculus modulo theory.From a practical and a theoretical viewpoint*, Université Paris Diderot (Paris 7), September 2017, https://hal.inria.fr/hal-01676409.

[25] F. GILBERT.*Extending higher-order logic with predicate subtyping: application to PVS*, December 2017, working paper or preprint, https://hal.inria.fr/hal-01673518.

[26] F. THIRÉ.*Exporting an Arithmetic Library from Dedukti to HOL*, December 2017, working paper or preprint, https://hal.inria.fr/hal-01668250.

References in notes

[27] Y. BERTOT, P. CASTÉRAN.*Interactive Theorem Proving and Program Development Coq'Art: The Calculus of Inductive Constructions*, Springer-Verlag, 2004.

[28] F. BOBOT, J.-C. FILLIÂTRE, C. MARCHÉ, A. PASKEVICH.*Why3: Shepherd Your Herd of Provers*, in "First International Workshop on Intermediate Verification Languages", 2011, http://hal.inria.fr/hal-00790310.

[29] M. BOESPFLUG.*Conception d'un noyau de vérification de preuves pour le lambda-Pi-calcul modulo*, École Polytechnique, 2011.

[30] G. DOWEK.*A Theory Independent Curry-de Bruijn-howard Correspondence*, in "Proceedings of the 39th International Colloquium Conference on Automata, Languages, and Programming - Volume Part II", Berlin, Heidelberg, ICALP'12, Springer-Verlag, 2012, p. 13–15, http://dx.doi.org/10.1007/978-3-642-31585-5_2.

[31] R. HARPER, F. HONSELL, G. PLOTKIN.*A Framework for Defining Logics*, in "Journal of the association for computing machinery", 1993, p. 194–204.

[32] J. HARRISON.*HOL Light: An Overview*, in "Theorem Proving in Higher Order Logics", S. BERGHOFER, T. NIPKOW, C. URBAN, M. WENZEL (editors), Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2009, vol. 5674, p. 60-66, http://dx.doi.org/10.1007/978-3-642-03359-9_4.

[33] J. HUGHES, L. PARETO, A. SABRY.*Proving the correctness of reactive systems using sized types*, in "Proceedings of the 23th ACM Symposium on Principles of Programming Languages", 1996, http://doi.org/10.1145/237721.240882.

[34] C. S. LEE, N. D. JONES, A. M. BEN-AMRAM.*The size-change principle for program termination*, in "Proceedings of the 28th ACM Symposium on Principles of Programming Languages", 2001.

[35] R. LEPIGRE.*PML_2, programming language with support for program proofs*, 2017, Submitted for the post-proceedings of TYPES, https://lepigre.fr/files/docs/lepigre2017_pml2.pdf.

[36] R. LEPIGRE.*Lambdapi, implementation of the $\lambda\Pi$-calculus modulo rewriting*, 2017, https://github.com/rlepigre/lambdapi.

[37] R. LEPIGRE.*Semantics and Implementation of an Extension of ML for program proving*, Université Grenoble Alpes, France, 2017, https://github.com/rlepigre/phd/blob/master/manuscript_archived.pdf.

[38] R. LEPIGRE, C. RAFFALLI.*Bindlib, representation of binders in OCaml*, 2015, https://github.com/rlepigre/ocaml-bindlib.

[39] R. LEPIGRE, C. RAFFALLI.*Practical Subtyping for System F with Sized (Co-)Induction*, 2017, Submitted to the TOPLAS journal (under revision), http://lepigre.fr/files/docs/subtyping2017.pdf.

[40] D. WAHLSTEDT.*Dependent Type Theory with Parameterized First-Order Data Types and Well-Founded Recursion*, Chalmers University of Technology, 2007, ISBN 978-91-7291-979-2, http://www.cse.chalmers.se/alumni/davidw/wdt_phd_printed_version.pdf.

Project-Team DEFI

Shape reconstruction and identification

IN COLLABORATION WITH: Centre de Mathématiques Appliquées (CMAP)

IN PARTNERSHIP WITH:
CNRS

Ecole Polytechnique

RESEARCH CENTER
Saclay - Île-de-France

THEME
Numerical schemes and simulations

Table of contents

Project-Team DEFI

Creation of the Project-Team: 2009 January 01

Keywords:

Computer Science and Digital Science:

A6. - Modeling, simulation and control
A6.1. - Mathematical Modeling
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.2. - Scientific Computing, Numerical Analysis & Optimization
A6.2.1. - Numerical analysis of PDE and ODE
A6.2.6. - Optimization
A6.3. - Computation-data interaction
A6.3.1. - Inverse problems
A6.3.5. - Uncertainty Quantification

Other Research Topics and Application Domains:

B1.2.1. - Understanding and simulation of the brain and the nervous system
B2.6.1. - Brain imaging
B3.3.1. - Earth and subsoil
B5.3. - Nanotechnology

1. Personnel

Research Scientists

Houssem Haddar [Team leader, Inria, Senior Researcher]
Lucas Chesnel [Inria, Researcher]
Jing-Rebecca Li [Inria, Researcher, HDR]

Faculty Member

Grégoire Allaire [Ecole polytechnique, Professor, HDR]

External Collaborators

Semra Ahmetolan [ITU, until Jan 2017]
Lorenzo Audibert [EDF]
Hanen Boujlida [ESSTHS, from Sep 2017 until Oct 2017]

PhD Students

Hugo Girardon [EDF, from Oct 2017]
Jingjing Hao [Inria, from Oct 2017]
Mohamed Lakhal [Ecole polytechnique, until Jun 2017]
Kevish Napal [Ecole polytechnique]
Van Khieu Nguyen [Univ Paris-Sud, until Mar 2017]

Post-Doctoral Fellows

Marc Bakry [Inria]
Imen Mekkaoui [Inria, from Nov 2017]
Thi Phong Nguyen [Inria, from Mar 2017]

Administrative Assistants

Natalia Alves [Inria]

Jessica Gameiro [Inria]
Maeva Jeannot [Inria]

2. Overall Objectives

2.1. Overall Objectives

The research activity of our team is dedicated to the design, analysis and implementation of efficient numerical methods to solve inverse and shape/topological optimization problems in connection with acoustics, electromagnetism, elastodynamics, and diffusion.

Sought practical applications include radar and sonar applications, bio-medical imaging techniques, non-destructive testing, structural design, composite materials, and diffusion magnetic resonance imaging.

Roughly speaking, the model problem consists in determining information on, or optimizing the geometry (topology) and the physical properties of unknown targets from given constraints or measurements, for instance, measurements of diffracted waves or induced magnetic fields.

In general this kind of problems is non-linear. The inverse ones are also severely ill-posed and therefore require special attention from regularization point of view, and non-trivial adaptations of classical optimization methods.

Our scientific research interests are the following:

- Theoretical understanding and analysis of the forward and inverse mathematical models, including in particular the development of simplified models for adequate asymptotic configurations.
- The design of efficient numerical optimization/inversion methods which are quick and robust with respect to noise. Special attention will be paid to algorithms capable of treating large scale problems (e.g. 3-D problems) and/or suited for real-time imaging.
- Development of prototype softwares for specific applications or tutorial toolboxes.

We were particularly interested in the development of the following themes

- Qualitative methods for inverse scattering problems
- Iterative and Hybrid inversion methods
- Topological optimization methods
- Forward and inverse models for Diffusion MRI
- Asymptotic models and methods for waves and diffusion.

3. Research Program

3.1. Research Program

The research activity of our team is dedicated to the design, analysis and implementation of efficient numerical methods to solve inverse and shape/topological optimization problems in connection with wave imaging, structural design, non-destructive testing and medical imaging modalities. We are particularly interested in the development of fast methods that are suited for real-time applications and/or large scale problems. These goals require to work on both the physical and the mathematical models involved and indeed a solid expertise in related numerical algorithms.

This section intends to give a general overview of our research interests and themes. We choose to present them through the specific academic example of inverse scattering problems (from inhomogeneities), which is representative of foreseen developments on both inversion and (topological) optimization methods. The practical problem would be to identify an inclusion from measurements of diffracted waves that result from the interaction of the sought inclusion with some (incident) waves sent into the probed medium. Typical applications include biomedical imaging where using micro-waves one would like to probe the presence of pathological cells, or imaging of urban infrastructures where using ground penetrating radars (GPR) one is interested in finding the location of buried facilities such as pipelines or waste deposits. This kind of applications requires in particular fast and reliable algorithms.

By “imaging” we refer to the inverse problem where the concern is only the location and the shape of the inclusion, while “identification” may also indicate getting informations on the inclusion physical parameters.

Both problems (imaging and identification) are non linear and ill-posed (lack of stability with respect to measurements errors if some careful constrains are not added). Moreover, the unique determination of the geometry or the coefficients is not guaranteed in general if sufficient measurements are not available. As an example, in the case of anisotropic inclusions, one can show that an appropriate set of data uniquely determine the geometry but not the material properties.

These theoretical considerations (uniqueness, stability) are not only important in understanding the mathematical properties of the inverse problem, but also guide the choice of appropriate numerical strategies (which information can be stably reconstructed) and also the design of appropriate regularization techniques. Moreover, uniqueness proofs are in general constructive proofs, i.e. they implicitly contain a numerical algorithm to solve the inverse problem, hence their importance for practical applications. The sampling methods introduced below are one example of such algorithms.

A large part of our research activity is dedicated to numerical methods applied to the first type of inverse problems, where only the geometrical information is sought. In its general setting the inverse problem is very challenging and no method can provide universally satisfying solution (respecting the balance cost-precision-stability). This is why in the majority of the practically employed algorithms, some simplification of the underlying mathematical model is used, according to the specific configuration of the imaging experiment. The most popular ones are geometric optics (the Kirchhoff approximation) for high frequencies and weak scattering (the Born approximation) for small contrasts or small obstacles. They actually give full satisfaction for a wide range of applications as attested by the large success of existing imaging devices (radar, sonar, ultrasound, X-ray tomography, etc.), that rely on one of these approximations.

In most cases, the used simplification result in a linearization of the inverse problem and therefore is usually valid only if the latter is weakly non-linear. The development of simplified models and the improvement of their efficiency is still a very active research area. With that perspective, we are particularly interested in deriving and studying higher order asymptotic models associated with small geometrical parameters such as: small obstacles, thin coatings, wires, periodic media, ... Higher order models usually introduce some non linearity in the inverse problem, but are in principle easier to handle from the numerical point of view than in the case of the exact model.

A larger part of our research activity is dedicated to algorithms that avoid the use of such approximations and that are efficient where classical approaches fail: i.e. roughly speaking when the non linearity of the inverse problem is sufficiently strong. This type of configuration is motivated by the applications mentioned below, and occurs as soon as the geometry of the unknown media generates non negligible multiple scattering effects (multiply-connected and closely spaces obstacles) or when the used frequency is in the so-called resonant region (wave-length comparable to the size of the sought medium). It is therefore much more difficult to deal with and requires new approaches. Our ideas to tackle this problem is mainly motivated and inspired by recent advances in shape and topological optimization methods and in so-called sampling methods.

Sampling methods are fast imaging solvers adapted to multi-static data (multiple receiver-transmitter pairs) at a fixed frequency. Even if they do not use any linearization the forward model, they rely on computing the solutions to a set of linear problems of small size, that can be performed in a completely parallel procedure.

Our team has already a solid expertise in these methods applied to electromagnetic 3-D problems. The success of such approaches was their ability to provide a relatively quick algorithm for solving 3-D problems without any need for a priori knowledge on the physical parameters of the targets. These algorithms solve only the imaging problem, in the sense that only the geometrical information is provided.

Despite the large efforts already spent in the development of this type of methods, either from the algorithmic point of view or the theoretical one, numerous questions are still open. These attractive new algorithms also suffer from the lack of experimental validations, due to their relatively recent introduction. We also would like to invest on this side by developing collaborations with engineering research groups that have experimental facilities. From the practical point of view, the most potential limitation of sampling methods would be the need of a large amount of data to achieve a reasonable accuracy. On the other hand, optimization methods do not suffer from this constrain but they require good initial guess to ensure convergence and reduce the number of iterations. Therefore it seems natural to try to combine the two class of methods in order to calibrate the balance between cost and precision.

Among various shape optimization methods, the Level Set method seems to be particularly suited for such a coupling. First, because it shares similar mechanism as sampling methods: the geometry is captured as a level set of an "indicator function" computed on a cartesian grid. Second, because the two methods do not require any a priori knowledge on the topology of the sought geometry. Beyond the choice of a particular method, the main question would be to define in which way the coupling can be achieved. Obvious strategies consist in using one method to pre-process (initialization) or post-process (find the level set) the other. But one can also think of more elaborate ones, where for instance a sampling method can be used to optimize the choice of the incident wave at each iteration step. The latter point is closely related to the design of so called "focusing incident waves" (which are for instance the basis of applications of the time-reversal principle). In the frequency regime, these incident waves can be constructed from the eigenvalue decomposition of the data operator used by sampling methods. The theoretical and numerical investigations of these aspects are still not completely understood for electromagnetic or elastodynamic problems.

Other topological optimization methods, like the homogenization method or the topological gradient method, can also be used, each one provides particular advantages in specific configurations. It is evident that the development of these methods is very suited to inverse problems and provide substantial advantage compared to classical shape optimization methods based on boundary variation. Their applications to inverse problems has not been fully investigated. The efficiency of these optimization methods can also be increased for adequate asymptotic configurations. For instance small amplitude homogenization method can be used as an efficient relaxation method for the inverse problem in the presence of small contrasts. On the other hand, the topological gradient method has shown to perform well in localizing small inclusions with only one iteration.

A broader perspective would be the extension of the above mentioned techniques to time-dependent cases. Taking into account data in time domain is important for many practical applications, such as imaging in cluttered media, the design of absorbing coatings or also crash worthiness in the case of structural design.

For the identification problem, one would like to also have information on the physical properties of the targets. Of course optimization methods is a tool of choice for these problems. However, in some applications only a qualitative information is needed and obtaining it in a cheaper way can be performed using asymptotic theories combined with sampling methods. We also refer here to the use of so called transmission eigenvalues as qualitative indicators for non destructive testing of dielectrics.

We are also interested in parameter identification problems arising in diffusion-type problems. Our research here is mostly motivated by applications to the imaging of biological tissues with the technique of Diffusion Magnetic Resonance Imaging (DMRI). Roughly speaking DMRI gives a measure of the average distance travelled by water molecules in a certain medium and can give useful information on cellular structure and structural change when the medium is biological tissue. In particular, we would like to infer from DMRI measurements changes in the cellular volume fraction occurring upon various physiological or pathological conditions as well as the average cell size in the case of tumor imaging. The main challenges here are 1) correctly model measured signals using diffusive-type time-dependent PDEs 2) numerically handle the complexity of the tissues 3) use the first two to identify physically relevant parameters from measurements.

For the last point we are particularly interested in constructing reduced models of the multiple-compartment Bloch-Torrey partial differential equation using homogenization methods.

4. Application Domains

4.1. Radar and GPR applications

Conventional radar imaging techniques (ISAR, GPR, etc.) use backscattering data to image targets. The commonly used inversion algorithms are mainly based on the use of weak scattering approximations such as the Born or Kirchhoff approximation leading to very simple linear models, but at the expense of ignoring multiple scattering and polarization effects. The success of such an approach is evident in the wide use of synthetic aperture radar techniques.

However, the use of backscattering data makes 3-D imaging a very challenging problem (it is not even well understood theoretically) and as pointed out by Brett Borden in the context of airborne radar: “In recent years it has become quite apparent that the problems associated with radar target identification efforts will not vanish with the development of more sensitive radar receivers or increased signal-to-noise levels. In addition it has (slowly) been realized that greater amounts of data - or even additional “kinds” of radar data, such as added polarization or greatly extended bandwidth - will all suffer from the same basic limitations affiliated with incorrect model assumptions. Moreover, in the face of these problems it is important to ask how (and if) the complications associated with radar based automatic target recognition can be surmounted.” This comment also applies to the more complex GPR problem.

Our research themes will incorporate the development, analysis and testing of several novel methods, such as sampling methods, level set methods or topological gradient methods, for ground penetrating radar application (imaging of urban infrastructures, landmines detection, underground waste deposits monitoring,) using multistatic data.

4.2. Biomedical imaging

Among emerging medical imaging techniques we are particularly interested in those using low to moderate frequency regimes. These include Microwave Tomography, Electrical Impedance Tomography and also the closely related Optical Tomography technique. They all have the advantage of being potentially safe and relatively cheap modalities and can also be used in complementarity with well established techniques such as X-ray computed tomography or Magnetic Resonance Imaging.

With these modalities tissues are differentiated and, consequentially can be imaged, based on differences in dielectric properties (some recent studies have proved that dielectric properties of biological tissues can be a strong indicator of the tissues functional and pathological conditions, for instance, tissue blood content, ischemia, infarction, hypoxia, malignancies, edema and others). The main challenge for these functionalities is to built a 3-D imaging algorithm capable of treating multi-static measurements to provide real-time images with highest (reasonably) expected resolutions and in a sufficiently robust way.

Another important biomedical application is brain imaging. We are for instance interested in the use of EEG and MEG techniques as complementary tools to MRI. They are applied for instance to localize epileptic centers or active zones (functional imaging). Here the problem is different and consists into performing passive imaging: the epileptic centers act as electrical sources and imaging is performed from measurements of induced currents. Incorporating the structure of the skull is primordial in improving the resolution of the imaging procedure. Doing this in a reasonably quick manner is still an active research area, and the use of asymptotic models would offer a promising solution to fix this issue.

4.3. Non destructive testing and parameter identification

One challenging problem in this vast area is the identification and imaging of defaults in anisotropic media. For instance this problem is of great importance in aeronautic constructions due to the growing use of composite materials. It also arises in applications linked with the evaluation of wood quality, like locating knots in timber in order to optimize timber-cutting in sawmills, or evaluating wood integrity before cutting trees. The anisotropy of the propagative media renders the analysis of diffracted waves more complex since one cannot only relies on the use of backscattered waves. Another difficulty comes from the fact that the micro-structure of the media is generally not well known a priori.

Our concern will be focused on the determination of qualitative information on the size of defaults and their physical properties rather than a complete imaging which for anisotropic media is in general impossible. For instance, in the case of homogeneous background, one can link the size of the inclusion and the index of refraction to the first eigenvalue of so-called interior transmission problem. These eigenvalues can be determined form the measured data and a rough localization of the default. Our goal is to extend this kind of idea to the cases where both the propagative media and the inclusion are anisotropic. The generalization to the case of cracks or screens has also to be investigated.

In the context of nuclear waste management many studies are conducted on the possibility of storing waste in a deep geological clay layer. To assess the reliability of such a storage without leakage it is necessary to have a precise knowledge of the porous media parameters (porosity, tortuosity, permeability, etc.). The large range of space and time scales involved in this process requires a high degree of precision as well as tight bounds on the uncertainties. Many physical experiments are conducted in situ which are designed for providing data for parameters identification. For example, the determination of the damaged zone (caused by excavation) around the repository area is of paramount importance since microcracks yield drastic changes in the permeability. Level set methods are a tool of choice for characterizing this damaged zone.

4.4. Diffusion MRI

In biological tissues, water is abundant and magnetic resonance imaging (MRI) exploits the magnetic property of the nucleus of the water proton. The imaging contrast (the variations in the grayscale in an image) in standard MRI can be from either proton density, T1 (spin-lattice) relaxation, or T2 (spin-spin) relaxation and the contrast in the image gives some information on the physiological properties of the biological tissue at different physical locations of the sample. The resolution of MRI is on the order of millimeters: the greyscale value shown in the imaging pixel represents the volume-averaged value taken over all the physical locations contained that pixel.

In diffusion MRI, the image contrast comes from a measure of the average distance the water molecules have moved (diffused) during a certain amount of time. The Pulsed Gradient Spin Echo (PGSE) sequence is a commonly used sequence of applied magnetic fields to encode the diffusion of water protons. The term 'pulsed' means that the magnetic fields are short in duration, an the term gradient means that the magnetic fields vary linearly in space along a particular direction. First, the water protons in tissue are labelled with nuclear spin at a precession frequency that varies as a function of the physical positions of the water molecules via the application of a pulsed (short in duration, lasting on the order of ten milliseconds) magnetic field. Because the precessing frequencies of the water molecules vary, the signal, which measures the aggregate phase of the water molecules, will be reduced due to phase cancellations. Some time (usually tens of milliseconds) after the first pulsed magnetic field, another pulsed magnetic field is applied to reverse the spins of the water molecules. The time between the applications of two pulsed magnetic fields is called the 'diffusion time'. If the water molecules have not moved during the diffusion time, the phase dispersion will be reversed, hence the signal loss will also be reversed, the signal is called refocused. However, if the molecules have moved during the diffusion time, the refocusing will be incomplete and the signal detected by the MRI scanner if weaker than if the water molecules have not moved. This lack of complete refocusing is called the signal attenuation and is the basis of the image contrast in DMRI. the pixels showning more signal attenuation is associated with further water displacement during the diffusion time, which may be linked to physiological factors, such as higher cell membrane permeability, larger cell sizes, higher extra-cellular volume fraction.

We model the nuclear magnetization of water protons in a sample due to diffusion-encoding magnetic fields by a multiple compartment Bloch-Torrey partial differential equation, which is a diffusive-type time-dependent PDE. The DMRI signal is the integral of the solution of the Bloch-Torrey PDE. In a homogeneous medium, the intrinsic diffusion coeffcient D will appear as the slope of the semi-log plot of the signal (in approporiate units). However, because during typical scanning times, 50-100ms, water molecules have had time to travel a diffusion distance which is long compared to the average size of the cells, the slope of the semi-log plot of the signal is in fact a measure of an 'effective' diffusion coefficient. In DMRI applications, this measured quantity is called the 'apparent diffusion coefficient' (ADC) and provides the most commonly used form the image contrast for DMRI. This ADC is closely related to the effective diffusion coefficient obtainable from mathematical homogenization theory.

5. New Software and Platforms

5.1. FVforBlochTorrey

- Participant: Jing Rebecca Li
- Contact: Jing Rebecca Li

5.2. InvGIBC

- Participant: Nicolas Chaulet
- Contact: Houssem Haddar

5.3. RODIN

FUNCTIONAL DESCRIPTION: In the framework of the RODIN project we continue to develop with our software partner ESI the codes Topolev and Geolev for topology and geometry shape optimization of mechanical structures using the level set method.

- Contact: Grégoire Allaire

5.4. samplings-2d

This software solves forward and inverse problems for the Helmholtz equation in 2-D.

FUNCTIONAL DESCRIPTION: This software is written in Fortran 90 and is related to forward and inverse problems for the Helmholtz equation in 2-D. It includes three independent components. * The first one solves to scattering problem using integral equation approach and supports piecewise-constant dielectrics and obstacles with impedance boundary conditions. * The second one contains various samplings methods to solve the inverse scattering problem (LSM, RGLSM(s), Factorization, MuSiC) for near-field or far-field setting. * The third component is a set of post processing functionalities to visualize the results

- Participant: Houssem Haddar
- Contact: Houssem Haddar
- URL: http://sourceforge.net/projects/samplings-2d/

5.5. Samplings-3d

FUNCTIONAL DESCRIPTION: This software is written in Fortran 90 and is related to forward and inverse problems for the Helmholtz equation in 3-D. It contains equivalent functionalities to samplings-2d in a 3-D setting.

- Contact: Houssem Haddar

5.6. SCILAB

SCIENTIFIC DESCRIPTION: Scilab includes hundreds of mathematical functions. It has a high level programming language allowing access to advanced data structures, 2-D and 3-D graphical functions.

A large number of functionalities is included in Scilab:

Maths & Simulation For usual engineering and science applications including mathematical operations and data analysis. 2-D & 3-D Visualization Graphics functions to visualize, annotate and export data and many ways to create and customize various types of plots and charts. Optimization Algorithms to solve constrained and unconstrained continuous and discrete optimization problems. Statistics Tools to perform data analysis and modeling Control System Design & Analysis Standard algorithms and tools for control system study Signal Processing Visualize, analyze and filter signals in time and frequency domains. Application Development Increase Scilab native functionalities and manage data exchanges with external tools. Xcos - Hybrid dynamic systems modeler and simulator Modeling mechanical systems, hydraulic circuits, control systems...
FUNCTIONAL DESCRIPTION: Scilab is free and open source software for numerical computation providing a powerful computing environment for engineering and scientific applications.

- Participant: Grégoire Allaire
- Contact: Grégoire Allaire
- URL: http://www.scilab.org/

6. New Results

6.1. Methods for inverse problems

6.1.1. The Generalized Linear Sampling Method for limited aperture measurements

L. Audibert and H. Haddar

We extend the so-called Generalized Linear Sampling Method (GLSM) to the case of limited aperture data at a fixed frequency. In this case the factorization of the sampling operator does not obey the symmetry required in the justification of the GLSM introduced in Audibert-Haddar [Inverse Problems, 2014]. We propose a new formulation by adding an extra penalty term that asymptotically correct the non symmetry of the GLSM original penalty term. The analysis of the new formulation is first presented in an abstract framework. We then show how to apply our setting to the scalar problem with far field measurements or near field measurements on a limited aperture. We finally validate the method through some numerical tests in two dimensions and for far field measurements.

6.1.2. A synoptic approach to the seismic sensing of heterogeneous fractures: from geometric reconstruction to interfacial characterization

B. Guzina, H. Haddar and F. Pourahmadian

A non-iterative waveform sensing approach is proposed toward (i) geometric reconstruction of penetrable fractures, and (ii) quantitative identification of their heterogeneous contact condition by seismic i.e. elastic waves. To this end, the fracture support Γ (which may be non-planar and unconnected) is first recovered without prior knowledge of the interfacial condition by way of the recently established approaches to non-iterative waveform tomography of heterogeneous fractures, e.g. the methods of generalized linear sampling and topological sensitivity. Given suitable approximation $\widetilde{\Gamma}$ of the fracture geometry, the jump in the displacement field across $\widetilde{\Gamma}$ i.e. the fracture opening displacement (FOD) profile is computed from remote sensory data via a regularized inversion of the boundary integral representation mapping the FOD to remote observations of the scattered field. Thus obtained FOD is then used as input for solving the traction boundary integral equation on $\widetilde{\Gamma}$ for the unknown (linearized) contact parameters. In this study, linear and possibly dissipative interactions between the two faces of a fracture are parameterized in terms of a symmetric, complex-valued matrix K collecting the normal, shear, and mixed-mode coefficients of specific stiffness. To facilitate the high-fidelity inversion for K, a 3-step regularization algorithm is devised to minimize the errors stemming from the inexact geometric reconstruction and FOD recovery. The performance of the inverse solution is illustrated by a set of numerical experiments where a cylindrical fracture, endowed with two example patterns of specific stiffness coefficients, is illuminated by plane waves and reconstructed in terms of its geometry and heterogeneous (dissipative) contact condition.

6.1.3. Sampling methods for reconstructing the geometry of a local perturbation in unknown periodic layers

H. Haddar and T.P Nguyen

The aim of this work is the design and analysis of sampling methods to reconstruct the shape of a local perturbation in a periodic layer from measurements of scattered waves at a fixed frequency. We first introduce the model problem that corresponds with the semi-discretized version of the continous model with respect to the Floquet-Bloch variable. We then present the inverse problem setting where (propagative and evanescent) plane waves are used to illuminate the structure and measurements of the scattered wave at a parallel plane to the periodicity directions are performed. We introduce the near field operator and analyze two possible factorizations of this operator. We then establish sampling methods to identify the defect and the periodic background geometry from this operator measurement. We also show how one can recover the geometry of the background independently from the defect. We then introduce and analyze the single Floquet-Bloch mode measurement operators and show how one can exploit them to built an indicator function of the defect independently from the background geometry. Numerical validating results are provided for simple and complex backgrounds.

6.1.4. Nanoparticles volume determination from SAXS measurements

M. Bakry and H. Haddar

The aim of this work is to develop a fully automatic method for the reconstruction of the volume distribution of polydisperse non-interacting nanoparticles with identical shapes from Small Angle X-ray Scattering measurements. In the case of diluted systems we proposed a method that solves a maximum likelihood problem with a positivity constraint on the solution by means of an Expectation Maximization iterative scheme coupled with a robust stopping criterion. We prove that the stopping rule provides a regularization method according to an innovative notion of regularization specifically defined for inverse problems with Poisson data. Such a regularization, together with the positivity constraint results in high fidelity quantitative reconstructions of particle volume distributions making the method particularly effective in real applications. We tested the performance of the method on synthetic data in the case of uni- and bi-modal particle volume distributions. We extended the method to the case of dense solutions where the inverse problem becomes non linear. The developement of this research topic is ongoing under the framework of Saxsize.

6.1.5. Identification of small objects with near-field data in quasi-backscattering configurations

H. Haddar and M. Lakhal

We present a new sampling method for detecting targets (small inclusions or defects) immersed in a homogeneous medium in three-dimensional space, from measurements of acoustic scattered fields created by point source incident waves. We consider the harmonic regime and a data setting that corresponds with quasi-backscattering configuration: the data is collected by a set a receivers that are distributed on a segment centered at the source position and the device is swept along a path orthogonal to the receiver line. We assume that the aperture of the receivers is small compared with the distance to the targets. Considering the asymptotic form of the scattered field as the size of the targets goes to zero and the small aperture approximation, one is able to derive a special expression for the scattered field. In this expression a separation of the dependence of scattered field on the source location and the distance source-target is performed. This allows us to propose a sampling procedure that characterizes the targets location in terms of the range of a near-field operator constructed from available data. Our procedure is similar to the one proposed by Haddar-Rezac for far-field configurations. The reconstruction algorithm is based on the MUSIC (Multiple SIgnal Classification) algorithm.

6.2. Invisiblity and transmission eigenvalues

6.2.1. Trapped modes and reflectionless modes as eigenfunctions of the same spectral problem

A.-S. Bonnet-Ben Dhia, L. Chesnel and V. Pagneux

We consider the reflection-transmission problem in a waveguide with obstacle. At certain frequencies, for some incident waves, intensity is perfectly transmitted and the reflected field decays exponentially at infinity. We show that such reflectionless modes can be characterized as eigenfunctions of an original non-selfadjoint spectral problem. In order to select ingoing waves on one side of the obstacle and outgoing waves on the other side, we use complex scalings (or Perfectly Matched Layers) with imaginary parts of different signs. We prove that the real eigenvalues of the obtained spectrum correspond either to trapped modes (or bound states in the continuum) or to reflectionless modes. Interestingly, complex eigenvalues also contain useful information on weak reflection cases. When the geometry has certain symmetries, the new spectral problem enters the class of $\mathcal{PT}$-symmetric problems.

6.2.2. Transmission eigenvalues with artificial background for explicit material index identification

L. Audibert, L. Chesnel and H. Haddar

We are interested in the problem of retrieving information on the refractive index n of a penetrable inclusion embedded in a reference medium from farfield data associated with incident plane waves. Our approach relies on the use of transmission eigenvalues (TEs) that carry information on n and that can be determined from the knowledge of the farfield operator F. We explain how to modify F into a farfield operator $F^{\mathrm{a}} = F - \widetilde{F}$, where $\widetilde{F}$ is computed numerically, corresponding to well chosen artificial background and for which the associated TEs provide more accessible information on n.

6.2.3. Simple examples of perfectly invisible and trapped modes in waveguides

L. Chesnel and V. Pagneux

We consider the propagation of waves in a waveguide with Neumann boundary conditions. We work at low wavenumber focusing our attention on the monomode regime. We assume that the waveguide is symmetric with respect to an axis orthogonal to the longitudinal direction and is endowed with a branch of height L whose width coincides with the wavelength of the propagating modes. In this setting, tuning the parameter L, we prove the existence of simple geometries where the transmission coefficient is equal to one (perfect invisibility). We also show that these geometries, for possibly different values of L, support so called trapped modes (non zero solutions of finite energy of the homogeneous problem) associated with eigenvalues embedded in the continuous spectrum.

6.2.4. Invisibility and perfect reflectivity in waveguides with finite length branches

L. Chesnel, S.A. Nazarov and V. Pagneux

We consider a time-harmonic wave problem, appearing for example in water-waves theory, in acoustics or in electromagnetism, in a setting such that the analysis reduces to the study of a 2D waveguide problem with a Neumann boundary condition. The geometry is symmetric with respect to an axis orthogonal to the direction of propagation of waves. Moreover, the waveguide contains one branch of finite length. We analyse the behaviour of the complex scattering coefficients $\mathcal{R}$, $\mathcal{T}$ as the length of the branch increases and we exhibit situations where non reflectivity ($\mathcal{R} = 0$, $|\mathcal{T}| = 1$), perfect reflectivity ($|\mathcal{R}| = 1$, $\mathcal{T} = 0$) or perfect invisibility ($\mathcal{R} = 0$, $\mathcal{T} = 1$) hold. Numerical experiments illustrate the different results.

6.2.5. Invisibility in scattering theory

L. Chesnel, A.-S. Bonnet-Ben Dhia and S.A. Nazarov

We are interested in a time harmonic acoustic problem in a waveguide with locally perturbed sound hard walls. We consider a setting where an observer generates incident plane waves at $-\infty$ and probes the resulting scattered field at $-\infty$ and $+\infty$. Practically, this is equivalent to measure the reflection and transmission coefficients respectively denoted R and T. In a recent work, a technique has been proposed to construct waveguides with smooth walls such that $R = 0$ and $|T| = 1$ (non reflection). However the approach fails to ensure $T = 1$ (perfect transmission without phase shift). First we establish a result explaining this observation. More precisely, we prove that for wavenumbers smaller than a given bound $k_{☆}$ depending on the geometry, we cannot have $T = 1$ so that the observer can detect the presence of the defect if he/she is able to measure the phase at $+\infty$. In particular, if the perturbation is smooth and small (in amplitude and in width), $k_{☆}$ is very close to the threshold wavenumber. Then, in a second step, we change the point of view and, for a given wavenumber, working with singular perturbations of the domain, we show how to obtain $T = 1$. In this case, the scattered field is exponentially decaying both at $-\infty$ and $+\infty$. We implement numerically the method to provide examples of such undetectable defects.

6.2.6. New sets of eigenvalues in inverse scattering for inhomogeneous media and their determination from scattering data

F. Cakoni, H. Haddar and L. Audibert

We developed a general mathematical framework to determine interior eigenvalues from a knowledge of the modified far field operator associated with an unknown (anisotropic) inhomogeneity. The modified far field operator is obtained by subtracting from the measured far field operator the computed far field operator corresponding to a well-posed scattering problem depending on one (possibly complex) parameter. Injectivity of this modified far field operator is related to an appropriate eigenvalue problem whose eigenvalues can be determined from the scattering data, and thus can be used to obtain information about material properties of the unknown inhomogeneity. We discuss here two examples of such modification leading to a Steklov eigenvalue problem, and a new type of the transmission eigenvalue problem. We present some numerical examples demonstrating the viability of our method for determining the interior eigenvalues form far field data.

6.2.7. The Asymptotic of Transmission Eigenvalues for a Domain with a Thin Coating

H. Boujlida, H Haddar and M. Khenissi

We consider the transmission eigenvalue problem for a medium surrounded by a thin layer of inhomogeneous material with different refractive index. We derive explicit asymptotic expansion for the transmission eigenvalues with respect to the thickness of the thin layer. We prove error estimate for the asymptotic expansion up to order 1 for simple eigenvalues. This expansion can be used to obtain explicit expressions for constant index of refraction.

6.3. Shape and topology optimization

6.3.1. Structural optimization under overhang constraints imposed by additive manufacturing technologies

G. Allaire, C. Dapogny, R. Estevez, A. Faure and G. Michailidis.

This work addresses one of the major constraints imposed by additive manufacturing processes on shape optimization problems - that of overhangs, i.e. large regions hanging over void without sufficient support from the lower structure. After revisiting the ‘classical’ geometric criteria used in the literature, based on the angle between the structural boundary and the build direction, we propose a new mechanical constraint functional, which mimics the layer by layer construction process featured by additive manufacturing technologies, and thereby appeals to the physical origin of the difficulties caused by overhangs. This constraint, as well as some variants, are precisely defined; their shape derivatives are computed in the sense of Hadamard’s method, and numerical strategies are extensively discussed, in two and three space dimensions, to efficiently deal with the appearance of overhang features in the course of shape optimization processes.

6.3.2. Shape optimisation with the level set method for contact problems in linearised elasticity

G. Allaire, F. Jouve and A. Maury

This work is devoted to shape optimisation of contact problems in linearised elasticity, thanks to the level set method. We circumvent the shape non-differentiability, due to the contact boundary conditions, by using penalised and regularised versions of the mechanical problem. This approach is applied to five different contact models: the frictionless model, the Tresca model, the Coulomb model, the normal compliance model and the Norton-Hoff model. We consider two types of optimisation problems in our applications: first, we minimise volume under a compliance constraint, second, we optimise the normal force, with a volume constraint, which is useful to design compliant mechanisms. To illustrate the validity of the method, 2D and 3D examples are performed, the 3D examples being computed with an industrial software.

6.3.3. Elasto-plastic shape optimization using the level set method

G. Allaire, F. Jouve and A. Maury

This work is concerned with shape optimization of structures made of a material obeying Hencky's laws of plasticity, with the stress bound expressed by the von Mises effective stress. The ill-posedness of the model is circumvented by using two regularized versions of the mechanical problem. The first one is the classical Perzyna formulation which is regularized, the second one is a new regularized formulation proposed for the von Mises criterion. Shape gradients are calculated thanks to the adjoint method. The optimal shape is numerically computed by using the level set method. To illustrate the validity of the method, 2D examples are performed.

6.4. Numerical methods for wave problems

6.4.1. Finite element methods for eigenvalue problems with sign-changing coefficients

C. Carvalho, P. Ciarlet and L. Chesnel

We consider a class of eigenvalue problems involving coefficients changing sign on the domain of interest. We analyse the main spectral properties of these problems according to the features of the coefficients. Under some assumptions on the mesh, we study how one can use classical finite element methods to approximate the spectrum as well as the eigenfunctions while avoiding spurious modes. We also prove localisation results of the eigenfunctions for certain sets of coefficients.

6.4.2. Linearized Navier-Stokes equations for Aeroacoustics using Stabilized Finite Elements : Boundary Conditions and Industrial Application to Aft-Fan Noise Propagation.

A. Bissuel, G. Allaire, L. Daumas, S., Barré and F. Rey

A numerical method for solving the linearized Navier-Stokes equations is presented for aeroacoustic sound propagation problem. The Navier-Stokes equations are linearized in the frequency domain. The fan noise of jet engine is emitted nearly selectively on some frequencies, which depend on the rotation velocity of the fan. A frequency domain approach is highly suitable for this kind of problems, instead of a costly time-dependent simulation which can handle a large range of frequencies depending on the time step and the mesh. The calculations presented here were all made using Aether, a Navier- Stokes code which uses finite elements stabilized with SUPG (Streamline Upwind Galerkin). Automatic code differentiation was used to linearize this code. Entropy variables bring interesting mathematical properties to the numerical scheme, but also prevent the easy implementation of boundary conditions. For instance, the pressure is a non-linear combination of the entropy variables. Imposing a pressure variation needs a linearization of this relation which is detailed herein. The performance of different types of boundary conditions used to impose the acoustic pressure variation inside the engine is studied in detail. Finally, a very surprising effect of the SUPG scheme was to transform a homogeneous Dirichlet boundary condition on all variables to a transparent one which is able to let only outgoing waves pass through with no incoming wave. A one-dimensional toy model is given to explain how SUPG brings about this transformation.

We finally treated an industrial test case. The geometry of a model turbine from the Clean Sky European project was used for sound propagation of the fan exhaust noise of a jet engine. Computations on several modes with increasing complexities were done and the results compared to a boundary element method which served as a reference when no mean flow is present. Results of a computation with a mean flow are shown.

6.5. Diffusion MRI

J.R. Li, K. V. Nguyen and I. Mekkaoui

Diffusion Magnetic Resonance Imaging (DMRI) is a promising tool to obtain useful information on microscopic structure and has been extensively applied to biological tissues.

We obtained the following results.

- The Bloch-Torrey equation describes the evolution of the spin (usually water proton) magnetization under the influence of applied magnetic field gradients and is commonly used in numerical simulations for diffusion MRI and NMR. Microscopic heterogeneity inside the imaging voxel is modeled by interfaces in- side the simulation domain, where a discontinuity in the magnetization across the interfaces is produced via a permeability coefficient on the interfaces. To avoid having to simulate on a computational domain that is the size of an en- tire imaging voxel, which is often much larger than the scale of the microscopic heterogeneity as well as the mean spin diffusion displacement, smaller represen- tative volumes of the imaging medium can be used as the simulation domain. In this case, the exterior boundaries of a representative volume either must be far away from the initial positions of the spins or suitable boundary conditions must be found to allow the movement of spins across these exterior boundaries.

  Many approaches have been taken to solve the Bloch-Torrey equation but an efficient high performance computing framework is still missing. We present formulations of the interface as well as the exterior boundary conditions that are computationally efficient and suitable for arbitrary order finite elements and parallelization. In particular, the formulations use extended finite elements with weak enforcement of real (in the case of interior interfaces) and artificial (in the case of exterior boundaries) permeability conditions as well as operator splitting for the exterior boundary conditions. The method is straightforward to implement and it is available in the FEniCS for moderate-scale simulations and in the FEniCS-HPC for large-scale simulations.
- The nerve cells of the *Aplysia* are much larger than mammalian neurons. Using the *Aplysia* ganglia to study the relationship between the cellular structure and the diffusion MRI signal can potentially shed light on this relationship for more complex organisms. We measured the dMRI signal of chemically-fixed abdominal ganglia of the *Aplysia* at several diffusion times. At the diffusion times measured, the dMRI signal is mono-exponential and can be accurately represented by the parameter ADC.

  We analyzed the diffusion time-dependent ADC using a well-known analytical formula that is valid in the short diffusion time regime. We performed this analysis for the largest sized cells of the ganglia to satisfy the short diffusion time requirement. We noted that a naive application of the short time formula is not adequate because of the presence of the cell nucleus, making the effective cell size much smaller than the actual cell size.

  We went on to perform numerical simulation of the ADC for several cell types of the abdominal ganglia. To create the simulation geometries, for the largest cells, we segmented a high resolution T2-weighted images and incorporated a manually generated nucleus. For small cells and nerve cells, we created spherical and cylindrical geometrical domains that are consistent with known information about the cellular structures from the literature. Using the library of simulation results, we fitted for the intrinsic diffusivities of the small cells and the nerve cells.
- We participated in providing simulation results for the Parietal team in their work on sensing Spindle Neurons in the Insula with Multi-shell Diffusion MRI.

- We started a new direction in the simulation and modeling of heart diffusion MRI with the post-doc project of Imen Mekkaoui, funded by Inria-EPFL lab. The project is co-supervised with Jan Hesthaven, Chair of Computational Mathematics and Simulation Science (MCSS), EPFL.

6.6. Mathematical tools for Psychology

J. R. Li and J. Hao

This is the start of a collaborative effort between the Defi team and the mental health professionals at the centre hospitalier Sainte Anne and l'Université Paris Diderot.

- We started a new research direction in algorithm and software development for analysis and classification of EEG measurements during the administration of neuropsychological tests for AD/HD with the PhD project of Jingjing Hao, co-supervised with Dr. Hassan Rahioui, Chef du pôle psychiatrique du 7e arrondissement de Paris rattaché au centre hospitalier Sainte-Anne.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

- A CIFRE PhD thesis started January 2015 with Dassault Aviations. The student is M. Aloïs Bissuel who is working on "linearized Navier-Stokes equations for optimization, fluttering and aeroacoustic".
- A CIFRE PhD thesis started December 2015 with Safran Tech. The student is Mrs Perle Geoffroy who is working on "topology optimization by the homogenization method in the context of additive manufacturing".
- A CIFRE PhD thesis started April 2017 with Safran Tech. The student is M. Florian Feppon who is working on "topology optimization for a coupled thermal-fluid-structure system".
- A CIFRE PhD thesis started October 2017 with Renault. The student is Mrs Lalaina Rakotondrainibe who is working on "topology optimization of connections between mechanical parts".
- A CIFRE PhD thesis started November 2017 with EDF. The student is H. Girardon who is working on "level set method for eddy current non destructive testting".

7.2. Bilateral Grants with Industry

- The SOFIA project (SOlutions pour la Fabrication Industrielle Additive métallique) started in the summer of 2016. Its purpose is to make research in the field of metallic additive manufacturing. The industrial partners include Michelin, FMAS, ESI, Safran and others. The academic partners are different laboratories of CNRS, including CMAP at Ecole Polytechnique. The project is funded for 6 years by BPI (Banque Publique d'Investissement).
- G. Allaire is participating to the TOP project at IRT SystemX which started in February 2017. It is concerned with the development of a topology optimization platform with industrial partners (Renault, Safran, Airbus, ESI).
- FUI project Tandem. This three years project started in December 2012 and has been extended to September 2017 involves Bull-Amesys (coordinator), BOWEN (ERTE+SART), Ecole Polytechnique (CMAP), Inria, LEAT et VSM. It aims at constructing a radar system on a flying device capable of real-time imaging mines embedded in dry soils (up to 40 cm deep). We are in charge of numerical validation of the inverse simulator.
- FUI project Saxsize. This three years project started in October 2015 and involves Xenocs (coordinator), Inria (DEFI), Pyxalis, LNE, Cordouan and CEA. It is a followup of Nanolytix where a focus is put on SAXS quantifications of dense nanoparticle solutions.

8. Partnerships and Cooperations

8.1. International Initiatives

8.1.1. Participation in Other International Programs

8.1.1.1. International Initiatives

QUASI

Title: Qualitative Approaches to Scattering and Imaging

International Partner (Institution - Laboratory - Researcher):

University of Rutgers (United States) - Fioralba Cakoni

Duration: 2013 - 2017

Start year: 2013

We concentrate on the use of qualitative methods in acoustic and electromagnetic inverse scattering theory with applications to nondestructive evaluation of materials and medical imaging. In particular, we would like to address theoretical and numerical reconstruction techniques to solve the inverse scattering problems using either time harmonic or time dependent measurements of the scattered field. The main goal of research in this field is to not only detect but also identify geometric and physical properties of unknown objects in real time.

8.2. International Research Visitors

8.2.1. Visits of International Scientists

- Fioralba Cakoni (2 weeks)
- David Colton (1 week)
- Armin Lechleiter (1 week)
- Rainer Kress (1 week)

8.2.1.1. Internships

- Marwa Kchaou (ENIT) 6 months
- FatmeMustapha (EDF) 6 months
- DucVu (Inria) 3 months

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. Member of the Organizing Committees

- G. Allaire is a member of the "comité national" CNRS, section 41 (mathematics). He is a member of the board of the Gaspard Monge program on optimization (PGMO) at the Jacques Hadamard Mathematical Foundation. He is a board member of Institut Henri Poincaré (IHP). He is the chairman of the scientific council of IFPEN (French Petroleum Institute and New Energies).
- L. Chesnel co-organized the Journée de rentrée (2017) of the Centre de Mathématiques Appliquées of École Polytechnique
- L. Chesnel co-organize the seminar of the Centre de Mathématiques Appliquées of École Polytechnique and the joint seminar of the Inria teams Defi-M3DISIM-Poems.

- H. Haddar Co-organized of the third Franco-German Summer School “Inverse Problems and Imaging”, University of Bremen, September 18-22, 2017
- DEFI was a sponsor of the conference "Waves diffracted by Patrick Joly", Paris, 2017.
- J.R. Li is organizer of Ecole d’ete d’excellence for Chinese Master’s students funded by French Embassy in China, 2017.
- J.R. Li is member of Organizing Committee of SIAM Conference on Computational Science and Engineering, 2017

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

- J.R. Li is member of the SIAM Committee on Programs and Conferences 2017-2019
- J.R. Li is responsable for the Ecole Polytechnique part of the French-Vietnam Master Program in Applied Mathematics, 2017
- J.R. Li is reviewer for Millennium Science Initiative, a program of the Government of Chile, 2017.
- H. Haddar is member of the scientific committees of the conferences series TAMTAM, Picof and Waves

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- G. Allaire is member of the editorial board of
  - book series "Mathématiques et Applications" of SMAI and Springer,
  - ESAIM/COCV, Structural and Multidisciplinary Optimization,
  - Discrete and Continuous Dynamical Systems Series B,
  - Computational and Applied Mathematics,
  - Mathematical Models and Methods in Applied Sciences (M3AS),
  - Annali dell’Universita di Ferrara,
  - OGST (Oil and Gas Science and Technology),
  - Journal de l’Ecole Polytechnique - Mathématiques,
  - Journal of Optimization Theory and Applications.
- H. Haddar is
  - member the editorial advisory board of Inverse Problems
  - Associate Editor of the SIAM Journal on Scientific Computing
  - Guest editor of Computers and Mathematics with Applications for a special issue on "Numerical Methods for PDEs and Inverse Problems"

9.1.3.2. Reviewer - Reviewing Activities

The members of the team reviewed numerous papers for numerous international journals. Too many to make a list.

9.1.4. Invited Talks

- G. Allaire
  - "GAMM-Seminar on Microstructures", Dortmund (January 2017).
  - Workshop on "Shape, Images and Optimization", Münster (March 2017).
  - Interaction of Applied Mathematics and Mechanics Conference, IAMMC2017, Paris (May 2017).
  - Congrès CSMA, Giens (May 2017).
  - New trends in shape optimization, Vosges (May 2017).
  - WCSMO, Braunschweig (June 2017).
  - CEDYA, Cartagena (June 2017).
  - Waves diffracted by Patrick Joly, Gif-sur-Yvette (August 2017).
  - SIM-AM ECCOMAS conference, Münich (October 2017).
- L. Chesnel
  - Worshop Inverse plasmonic problems-Neumann Poincaré operator, Université de Grenoble-Alpes, November 2017.
  - Séminaire EDP, modélisation et calcul scientifique, UMPA, ENS Lyon, November 2017.
  - Séminaire EDP, analyse et applications, Université de Lorraine, Metz, November 2017.
  - Séminaire EDP/Physique mathématique, Université de Bordeaux, September 2017.
  - Waves diffracted by Patrick Joly, Gif-sur-Yvette, August 2017.
- H. Haddar
  - TamTam'17, Hammamet, Tunisia, May 2017, Minisymposium, Inverse and imaging problems for PDE with applications
  - Oberwolfach Workshop Computational Inverse Problems for Partial Differential Equations, Oberwolfach, May 2017.
  - Workshop on the occasion of the 75th birthday of Rainer Kress, Goettingen, May 2017
  - Applied Inverse Problems, Hangzou, May 2017
    - Minisymposium on Inverse Spectral Problems
    - Minisymposium on Stability and reconstruction in inverse problems and their applications
    - Minisymposium on Recent Developments on Computation of Transmission Eigenvalues with Applications
  - Workshop on nonlinear analysis : Recent advances and new trends, Monastir, July 2017
  - Quantitative Tomographic Imaging : Radon meets Bell and Maxwell, RICAM, Linz, July 2017
  - Waves diffracted by Patrick Joly, Gif-sur-Yvette, August 2017.
  - Colloquium of the mathematical department, Mainz University, October 2017.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

- Master : Grégoire Allaire, Approximation Numérique et Optimisation, for students in the second year of Ecole Polytechnique curriculum: 8 lessons of 1h30.
- Master : Houssem Haddar, Approximation Numérique et Optimisation, for students in the second year of Ecole Polytechnique curriculum: 8 TDs of 4h.

- Master : Houssem Haddar, Variational analysis of partial differential equations, for students in the second year of Ecole Polytechnique curriculum: 8 TDs of 4h.
- Master : Lucas Chesnel, “Variational analysis for partial differential equations”, 16 equivalent TD hours, second year (2A), École Polytechnique, Palaiseau, France
- Master : Lucas Chesnel, “Numerical approximation and optimisation”, 14 equivalent TD hours, second year (2A), École Polytechnique, Palaiseau, France
- Master : Lucas Chesnel, “Elementary tools of analysis for partial differential equations”, 25 equivalent TD hours, L3, Ensta ParisTech, Palaiseau, France
- Master : Grégoire Allaire, Optimal design of structures, for students in the third year of Ecole Polytechnique curriculum. 9 lessons of 1h30.
- Master : Grégoire Allaire, Theoretical and numerical analysis of hyperbolic systems of conservation laws, Master M2 ”mathematical modeling”, 8 lessons of 3h.
- Master : Jing Rebecca Li, Mathematical and numerical foundations of modeling and simulation using partial differential equations. French-Vietnam Master Program in Applied Mathematics.
- Doctorat : Houssem Haddar, Lecturer at the Summer School on Quantitative Tomographic Imaging : Radon meets Bell and Maxwell. (2x1h30) July 10-14, RICAM, Linz, 2017.

9.2.2. Supervision

- Ph.D. : M. Lakhal, Méthodes d’inversion pour la reconstruction de mines enfouies à partir de mesures d’antennes radar, June 2017, H. Haddar
- Ph.D. : T.P. Nguyen, Direct and inverse solvers for scattering problems from locally perturbed infinite periodic layers, January 2017, H. Haddar
- Ph.D. : K. Van Nguyen, Modeling, simulation and experimental verification of water diffusion in neuronal network of the Aplysia ganglia, March 2017, J.-R. Li and L. Ciobanu
- Ph.D. in progress: B. Charfi, Identification of the sigular support of a GIBC, 2014, H. Haddar and S. Chaabane
- PhD in progress : A. Talpaert, the direct numerical simulation of vapor bubbles at low Mach number with adaptative mesh refinement, 2013, G. Allaire and S. Dellacherie
- PhD in progress : A. Bissuel, linearized Navier Stokes equations for optimization, floating and aeroaccoustic, 2014, G. Allaire
- PhD in progress : P. Geoffroy on topology optimization by the homogenization method in the context of additive manufacturing (Safran Tech, to be defended in 2019), G. Allaire.
- PhD in progress : S. Houbar sur la cavitation dans le fluide caloporteur induite par les mouvements des assemblages d’un réacteur (CEA, to be defended in 2020), G. Allaire and G. Campioni
- PhD in progress : M. Boissier sur l’optimisation couplée de la topologie des formes et de la trajectoire de lasage en fabrication additive (to be defended in 2020). G. Allaire and Ch. Tournier.
- PhD in progress : L. Rakotondrainibe sur l’optimisation des liaisons enre pièces dans les système mécaniques (to be defended in 2020), G. Allaire.
- PhD in progress : F. Feppon sur l’optimisation topologique de systèmes couplés fluide-solide-thermique (Safran, to be defended in 2020), G. allaire and Ch. Dapogny.
- PhD in progress : Q. Feng sur les éléments finis multi-échelles pour Navier Stokes incompressible en milieu encombré (CEA, to be defended in 2019), G. Allaire and A. Cartalade.
- PhD in progress : K. Napal, Transmission eigenvalues and non destructive testing of concrete like materials , 2016, L. Chesnel H. Haddar and L. Audibert
- PhD in progress : M. Kchaou, Higher order homogenization tensors for DMRI modeling, 2016, H. Haddar, J.R Li and M. Moakher

- PhD in progress : H. Girardon, Non destructive testing of PWR tubes using eddy current rotating coils, 2017, H. Haddar and L. Audibert
- PhD in progress : J. Hao, Thesis topic: Algorithm and software development for analysis and classification of EEG measurements during administration of neuropsychological tests for AD/HD, 2017, J.R. Li and H. Rahioui.

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] L. AUDIBERT.*Sampling method for sign changing contrast*, in "Inverse Problems and Imaging ", 2017, https://hal.archives-ouvertes.fr/hal-01422024.

[2] L. AUDIBERT, F. CAKONI, H. HADDAR.*New sets of eigenvalues in inverse scattering for inhomogeneous media and their determination from scattering data*, in "Inverse Problems", December 2017, vol. 33, n^{o} 12, p. 1-30 [*DOI :* 10.1088/1361-6420/AA982F], https://hal.inria.fr/hal-01645862.

[3] L. AUDIBERT, H. HADDAR.*The Generalized Linear Sampling Method for limited aperture measurements*, in "SIAM Journal on Imaging Sciences", 2017, vol. 10, n^{o} 2, p. 845–870 [*DOI :* 10.1137/16M110112X], https://hal.archives-ouvertes.fr/hal-01422027.

[4] M. BAKRY, S. PERNET, F. COLLINO.*A new accurate residual-based a posteriori error indicator for the BEM in 2D-acoustics*, in "Computers and Mathematics with Applications", 2017, vol. 73, n^{o} 12, p. 2501 - 2514 [*DOI :* 10.1016/J.CAMWA.2017.03.016], https://hal.archives-ouvertes.fr/hal-01654558.

[5] A.-S. BONNET-BEN DHIA, L. CHESNEL, S. NAZAROV.*Perfect transmission invisibility for waveguides with sound hard walls*, in "Journal de Mathématiques Pures et Appliquées", August 2017, https://hal.archives-ouvertes.fr/hal-01371163.

[6] C. CARVALHO, L. CHESNEL, P. CIARLET.*Eigenvalue problems with sign-changing coefficients*, in "Comptes Rendus Mathématique", June 2017, vol. 355, n^{o} 6, p. 671 - 675 [*DOI :* 10.1016/J.CRMA.2017.05.002], https://hal.archives-ouvertes.fr/hal-01394856.

[7] L. A. CHESNEL, X. CLAEYS, S. A. NAZAROV.*Small obstacle asymptotics for a 2D semi-linear convex problem*, in "Applicable Analysis", February 2017, 20 [*DOI :* 10.1080/00036811.2017.1295449], https://hal.archives-ouvertes.fr/hal-01427617.

[8] G. DURRIEU, E. FRÉNOD, T. MORINEAU, T. NGUYEN.*Modeling Abstraction Hierarchy Levels of the Cyber Attacks Using Random Process*, in "Open Journal of Statistics", 2017, vol. 07, n^{o} 03, p. 500 - 520 [*DOI :* 10.4236/OJS.2017.73035], https://hal.archives-ouvertes.fr/hal-01670520.

[9] M. GIACOMINI.*An equilibrated fluxes approach to the certified descent algorithm for shape optimization using conforming finite element and discontinuous Galerkin discretizations*, in "Journal of Scientific Computing", 2017, https://arxiv.org/abs/1611.03391v2 [*DOI :* 10.1007/S10915-017-0545-1], https://hal.archives-ouvertes.fr/hal-01395529.

[10] M. GIACOMINI, O. PANTZ, K. TRABELSI.*Certified Descent Algorithm for shape optimization driven by fully-computable a posteriori error estimators*, in "ESAIM: Control, Optimisation and Calculus of Variations", 2017, vol. 23, n° 3, p. 977-1001, https://arxiv.org/abs/1604.04921 [*DOI :* 10.1051/COCV/2016021], https://hal.archives-ouvertes.fr/hal-01201914.

[11] H. HADDAR, J.-R. LI, S. SCHIAVI.*Understanding the Time-Dependent Effective Diffusion Coefficient Measured by Diffusion MRI: the Intra-Cellular Case*, in "SIAM Journal on Applied Mathematics", 2017, https://hal.inria.fr/hal-01421928.

[12] H. HADDAR, T.-P. NGUYEN.*Sampling methods for reconstructing the geometry of a local perturbation in unknown periodic layers*, in "Computers and Mathematics with Applications", December 2017, vol. 74, n° 11, p. 2831-2855 [*DOI :* 10.1016/J.CAMWA.2017.07.015], https://hal.inria.fr/hal-01645674.

[13] F. POURAHMADIAN, B. B. GUZINA, H. HADDAR.*A synoptic approach to the seismic sensing of heterogeneous fractures: from geometric reconstruction to interfacial characterization*, in "Computer Methods in Applied Mechanics and Engineering", September 2017, vol. 324, p. 395-412 [*DOI :* 10.1016/J.CMA.2017.06.002], https://hal.inria.fr/hal-01422085.

[14] P. SVEHLA, K.-V. NGUYEN, J.-R. LI, L. CIOBANU.*Quantitative DLA-based compressed sensing for T1-weighted acquisitions*, in "Journal of Magnetic Resonance", 2017 [*DOI :* 10.1016/J.JMR.2017.05.002], https://hal.archives-ouvertes.fr/hal-01429506.

Books or Proceedings Editing

[15] H. HADDAR, F. CAKONI, J. SUN, C. BACUTA (editors). *Proceedings of the international conference on computational mathematics and inverse problems honoring Peter Monk*, Elsevier, France, December 2017 [*DOI :* 10.1016/J.CAMWA.2017.09.001], https://hal.inria.fr/hal-01645753.

Other Publications

[16] G. ALLAIRE, L. JAKABČIN.*Taking into account thermal residual stresses in topology optimization of structures built by additive manufacturing*, December 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01666081.

[17] L. AUDIBERT, L. A. CHESNEL, H. HADDAR.*Transmission eigenvalues with artificial background for explicit material index identification*, November 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01631062.

[18] H. BOUJLIDA, H. HADDAR, M. KHENISSI.*The Asymptotic of Transmission Eigenvalues for a Domain with a Thin Coating*, November 2017, working paper or preprint, https://hal.inria.fr/hal-01646003.

[19] L. A. CHESNEL, S. A. NAZAROV, V. PAGNEUX.*Invisibility and perfect reflectivity in waveguides with finite length branches*, September 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01469833.

[20] L. CHESNEL, V. PAGNEUX.*Simple examples of perfectly invisible and trapped modes in waveguides*, September 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01593226.

[21] M. GIACOMINI, O. PANTZ, K. TRABELSI.*Volumetric expressions of the shape gradient of the compliance in structural shape optimization*, January 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01441943.

[22] A. MAURY, G. ALLAIRE, F. JOUVE.*Shape optimisation with the level set method for contact problems in linearised elasticity*, January 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01435325.

Project-Team DISCO

Dynamical Interconnected Systems in COmplex Environments

IN COLLABORATION WITH: Laboratoire des signaux et systèmes (L2S)

IN PARTNERSHIP WITH:
CNRS

CentraleSupélec

RESEARCH CENTER
Saclay - Île-de-France

THEME
Optimization and control of dynamic systems

Table of contents

Project-Team DISCO

Creation of the Team: 2010 January 01, updated into Project-Team: 2012 January 01

Keywords:

Computer Science and Digital Science:

A3.4.4. - Optimization and learning
A3.4.5. - Bayesian methods
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.3. - Discrete Modeling (multi-agent, people centered)
A6.4.1. - Deterministic control
A6.4.3. - Observability and Controlability
A6.4.4. - Stability and Stabilization

Other Research Topics and Application Domains:

B2.2.3. - Cancer
B2.3. - Epidemiology
B3.6. - Ecology
B4.3.3. - Wind energy
B5.2.3. - Aviation
B7.2.1. - Smart vehicles

1. Personnel

Research Scientists

Catherine Bonnet [Team leader, Inria, Senior Researcher, HDR]
Frédéric Mazenc [Inria, Researcher, HDR]
Silviu-Iulian Niculescu [CNRS, Senior Researcher, HDR]

Faculty Members

Sorin Olaru [CentraleSupélec, Professor, HDR]
Guillaume Sandou [CentraleSupélec, Professor, HDR]
Ali Zemouche [Univ de Lorraine, Associate Professor, until Aug 2017, HDR]

External Collaborators

Islam Boussaada [IPSA/CNRS]
Ali Zemouche [Univ de Lorraine, from Sep 2017]

PhD Students

Saeed Ahmed [Inria]
Caetano Cardeliquio [UNICAMP & Inria]
Walid Djema [Inria, until Nov 2017]
Dina Irofti [CentraleSupélec, until Sep 2017]

Post-Doctoral Fellow

Le Ha Vy Nguyen [Inria]

Visiting Scientists

Stefanella Boatto [Univ Federale Rio de Janeiro, from Mar 2017]
André Fioravanti [UNICAMP, until Feb 2017]
Yutaka Yamamoto [Kyoto University, from May 2017 until Aug 2017]

Administrative Assistant
Katia Evrat [Inria]

2. Overall Objectives

2.1. Objectives

The goal of the project is to better understand and well formalize the effects of complex environments on the dynamics of the interconnections, as well as to develop new methods and techniques for the analysis and control of such systems.

It is well-known that the interconnection of dynamic systems has as consequence an increased complexity of the behavior of the total system.

In a simplified way, as the concept of dynamics is well-understood, the interconnections can be seen as associations (by connections of materials or information flows) of distinct systems to ensure a pooling of the resources with the aim of obtaining a better operation with the constraint of continuity of the service in the event of a fault. In this context, the environment can be seen as a collection of elements, structures or systems, natural or artificial constituting the neighborhood of a given system. The development of interactive games through communication networks, control from distance (e.g. remote surgical operations) or in hostile environment (e.g. robots, drones), as well as the current trend of large scale integration of distribution (and/or transport and/or decision) and open information systems with systems of production, lead to new modeling schemes in problems where the dynamics of the environment have to be taken into account.

In order to tackle the control problems arising in the above examples, the team investigates new theoretical methods, develops new algorithms and implementations dedicated to these techniques.

3. Research Program

3.1. Analysis of interconnected systems

The major questions considered are those of the characterization of the stability (also including the problems of sensitivity compared to the variations of the parameters) and the determination of stabilizing controllers of interconnected dynamic systems. In many situations, the dynamics of the interconnections can be naturally modelled by systems with delays (constant, distributed or time-varying delays) eventually of fractional order. In other cases, partial differential equations (PDE) models can be better represented or approximated by using systems with delays. Our expertise on this subject, on both time and frequency domain methods, allows us to challenge difficult problems (e.g. systems with an infinite number of unstable poles).

- Robust stability of linear systems

  Within an interconnection context, lots of phenomena are modelled directly or after an approximation by delay systems. These systems might have fixed delays, time-varying delays, distributed delays ...

  For various infinite-dimensional systems, particularly delay and fractional systems, input-output and time-domain methods are jointly developed in the team to characterize stability. This research is developed at four levels: analytic approaches (H_∞-stability, BIBO-stablity, robust stability, robustness metrics) [1], [2], [5], [6], symbolic computation approaches (SOS methods are used for determining easy-to-check conditions which guarantee that the poles of a given linear system are not in the closed right half-plane, certified CAD techniques), numerical approaches (root-loci, continuation methods) and by means of softwares developed in the team [5], [6].

- Robustness/fragility of biological systems

Deterministic biological models describing, for instance, species interactions, are frequently composed of equations with important disturbances and poorly known parameters. To evaluate the impact of the uncertainties, we use the techniques of designing of global strict Lyapunov functions or functional developed in the team.

However, for other biological systems, the notion of robustness may be different and this question is still in its infancy (see, e.g. [70]). Unlike engineering problems where a major issue is to maintain stability in the presence of disturbances, a main issue here is to maintain the system response in the presence of disturbances. For instance, a biological network is required to keep its functioning in case of a failure of one of the nodes in the network. The team, which has a strong expertise in robustness for engineering problems, aims at contributing at the develpment of new robustness metrics in this biological context.

3.2. Stabilization of interconnected systems

- Linear systems: Analytic and algebraic approaches are considered for infinite-dimensional linear systems studied within the input-output framework.

  In the recent years, the Youla-Kučera parametrization (which gives the set of all stabilizing controllers of a system in terms of its coprime factorizations) has been the cornerstone of the success of the H_∞-control since this parametrization allows one to rewrite the problem of finding the optimal stabilizing controllers for a certain norm such as H_∞ or H_2 as affine, and thus, convex problem.

  A central issue studied in the team is the computation of such factorizations for a given infinite-dimensional linear system as well as establishing the links between stabilizability of a system for a certain norm and the existence of coprime factorizations for this system. These questions are fundamental for robust stabilization problems [1], [2].

  We also consider simultaneous stabilization since it plays an important role in the study of reliable stabilization, i.e. in the design of controllers which stabilize a finite family of plants describing a system during normal operating conditions and various failed modes (e.g. loss of sensors or actuators, changes in operating points). Moreover, we investigate strongly stabilizable systems, namely systems which can be stabilized by stable controllers, since they have a good ability to track reference inputs and, in practice, engineers are reluctant to use unstable controllers especially when the system is stable.
- Nonlinear systems

  The project aims at developing robust stabilization theory and methods for important classes of nonlinear systems that ensure good controllerperformance under uncertainty and time delays. The main techniques include techniques called backstepping and forwarding, contructions of strict Lyapunov functions through so-called "strictification" approaches [3] and construction of Lyapunov-Krasovskii functionals [4], [5], [6].
- Predictive control

  For highly complex systems described in the time-domain and which are submitted to constraints, predictive control seems to be well-adapted. This model based control method (MPC: Model Predictive Control) is founded on the determination of an optimal control sequence over a receding horizon. Due to its formulation in the time-domain, it is an effective tool for handling constraints and uncertainties which can be explicitly taken into account in the synthesis procedure [7]. The team considers how mutiparametric optimization can help to reduce the computational load of this method, allowing its effective use on real world constrained problems.

  The team also investigates stochastic optimization methods such as genetic algorithm, particle swarm optimization or ant colony [8] as they can be used to optimize any criterion and constraint whatever their mathematical structure is. The developed methodologies can be used by non specialists.

3.3. Synthesis of reduced complexity controllers

- PID controllers

  Even though the synthesis of control laws of a given complexity is not a new problem, it is still open, even for finite-dimensional linear systems. Our purpose is to search for good families of “simple” (e.g. low order) controllers for infinite-dimensional dynamical systems. Within our approach, PID candidates are first considered in the team [2], [71].
- Predictive control

  The synthesis of predictive control laws is concerned with the solution of multiparametric optimization problems. Reduced order controller constraints can be viewed as non convex constraints in the synthesis procedure. Such constraints can be taken into account with stochastic algorithms.

Finally, the development of algorithms based on both symbolic computation and numerical methods, and their implementations in dedicated Scilab/Matlab/Maple toolboxes are important issues in the project.

4. Application Domains

4.1. Analysis and Control of life sciences systems

The team is involved in life sciences applications. The two main lines are the analysis of bioreactors models and the modeling of cell dynamics in Acute Myeloblastic Leukemias (AML) in collaboration with St Antoine Hospital in Paris. A recent new subject is the modelling of Dengue epidemics.

4.2. Energy Management

The team is interested in Energy management and considers optimization and control problems in energy networks.

5. Highlights of the Year

5.1. Highlights of the Year

Silviu-Iulian Niculescu is a 2018 IEEE Control Systems Society Fellow for research on the effects of delays in system dynamics.

6. New Software and Platforms

6.1. YALTA

Yet Another LTI TDS Algorithm
FUNCTIONAL DESCRIPTION: The YALTA toolbox is a Matlab toolbox dedicated to the study of classical and fractional systems with delay in the frequency-domain. Its objective is to provide basic but important information such as, for instance, the position of the neutral chains of poles and unstable poles, as well as the root locus with respect to the delay of the system. The corresponding algorithms are based on recent theoretical results and on classical continuation methods exploiting the particularities of the problem.

- Participants: André Fioravanti, Catherine Bonnet, David Avanessoff, Hugo Cavalera, Jim Pioche and Le Ha Vy Nguyen
- Contact: Catherine Bonnet
- URL: http://yalta-toolbox.gforge.inria.fr/

7. New Results

7.1. Maximal-multiplicity-based rightmost-root assignment for retarded TDS

Participants: Islam Boussaada, Silviu-Iulian Niculescu, Sami Tliba [L2S], Hakki Unal [Anadolu University], Toma Vyhlidal [Czech Technical University].

The proposed approach is a stabilizing delayed state-feedback design guaranteeing an appropriate (admissible) convergence rate to the trivial solution of the controlled dynamical system. Unlike methods based on finite spectrum assignment, our method does not render the closed loop system finite dimensional but consists in controlling its rightmost spectral value. First, it consists in characterizing the root of the characteristic quasipolynomial function to be of maximal multiplicity by mean of an analytical necessary and sufficient condition. Then, conditions on such a root (of maximal multiplicity) to be stable and dominant are established. These results are obtained for reduced-orders time-delay system (scalar and quadratic cases), see [69].

7.2. Migration of multiple roots under parameters/delays perturbation

Participants: Islam Boussaada, Dina Irofti, Silviu-Iulian Niculescu, Wim Michiels [KU Leuven].

In the context of the perturbation theory of nonlinear eigenvalue problem, the sensitivity of multiple eigenvalues with respect to parameters' variations is studied. In the complete regular splitting case, explicit expressions for the leading coefficients of the Puiseux series of the eigenvalue are provided [22]. In contrast to existing analysis of multiple roots of delay equations the developed results are in a matrix framework, i.e., without reduction of the problem to the analysis of a scalar characteristic quasipolynomial.

7.3. A generalized $\tau-$decomposition for TDS with delay-dependent coefficients

Participants: Chi Jin [L2S], Keqin Gu [Illinois State University], Islam Boussaada, Silviu-Iulian Niculescu.

The standard frequency domain approaches for Time-delay systems analysis do not apply when the coefficients of the system are delay-dependent. Given a system with delay-dependent coefficients as well as a delay interval of interest, a method is proposed to find all the delay subintervals guaranteeing the asymptotic stability of the trivial solution. The crossing direction criteria is proposed which can be clearly interpreted from a geometrical two-parameter perspective [36], [52].

7.4. State and Output-feedback control design for (possibly fractional) time-delay systems

Participants: Catherine Bonnet, Caetano Cardeliquio, André Fioravanti [FEM-UNICAMP, Brazil].

We obtained this year new results for H_∞-control synthesis via output-feedback through a finite order LTI system, called comparison system [42].

We also generalised those results for fractional systems.

The fractional comparison system was obtained and through LMIs we were able to calculate the H_∞-norm for the fractional system and design a state-feedback control through the comparison system approach.

7.5. Stability and Stabilisability Through Envelopes for Retarded and Neutral Time-Delay Systems

Participants: Catherine Bonnet, Caetano Cardeliquio, Silviu Niculescu, André Fioravanti [FEM-UNICAMP, Brazil].

We presented a new approach to develop an envelope that engulfs all poles of a time-delay system.

Through LMIs we determined envelopes for retarded and neutral time-delay systems.

The envelopes proposed were not only tighter than the ones in the literature but they can also be applied to verify the stability of the system.

The approach was also used to design state-feedback controllers which cope with design requirements regarding $\alpha - stability$.

7.6. Backstepping with artificial delays

Participants: Frederic Mazenc, Michael Malisoff [LSU, USA], Laurent Burlion [ONERA], Victor Gibert [Airbus], Jerome Weston [LSU, USA].

We worked on the problem of improving a fundamental control design technique called backstepping.

We provided in [54] a new backstepping control design for time-varying systems with input delays. The result was obtained by the introduction of a constant 'artificial' pointwise delay in the input and a dynamic extension. Thus it is significantly different from backstepping results for systems with delay in the input as presented in previous contributions. The result in [54] ensures global asymptotic convergence for a broad class of partially linear systems with an arbitrarily large number of integrators. We used only one artificial delay, and we assumed that the nonlinear subsystems satisfy a converging-input-converging-state assumption. When the nonlinear subsystem is control affine with the state of the first integrator as the control, we provided sufficient conditions for our converging-input- converging-state assumption to hold.

7.7. Stability of time-varying systems with delay and Switched Nonlinear Systems

Participants: Frederic Mazenc, Hitay Ozbay [Blikent University, Turkey], Saeed Ahmed [Blikent University, Turkey], Silviu Niculescu, Michael Malisoff [LSU, USA].

Switched systems is a family of systems which is frequently encountered in practice and can be used to approximate time-varying systems to ease their stability analysis or control. In the two works [20] and [17], we provided results that are useful when it comes to analyze the stability of time-varying or switched systems with delay. In [20] we provided several significant applications of the trajectory approach developed recently by Mazenc and Malisoff. In two results, we used a Lyapunov function for a corresponding undelayed system to provide a new method for proving stability of linear continuous-time time-varying systems with bounded time-varying delays. Our main results used upper bounds on an integral average involving the delay. We also provided a novel reduction model approach that ensures global exponential stabilization of linear systems with a time-varying pointwise delay in the input, which allows the delay to be discontinuous and uncertain.

Three of our other works are devoted to switched systems. In [55] and [21], a new technique is proposed to ensure global asymptotic stability for nonlinear switched time-varying systems with time-varying discontinuous delays. It uses an adaptation of Halanay's inequality to switched systems and the trajectory based technique mentioned above. The result is applied to a family of linear time-varying systems with time-varying delays. In [53], we presented an extension of the trajectory based approach mentioned above for state feedback stabilization of switched linear continuous-time systems with a time-varying input delay. In contrast with finding classical common Lyapunov function or multiple Lyapunov functions for establishing the stability of the closed-loop switched system, the new trajectory based approach relies on verifying certain inequalities along the solution of a supplementary system. This study does not make any assumption regarding the stabilizability of all of the constituent modes of the switched system. Moreover, no assumption is needed about the differentiability of the delay and no constraint is imposed on the upper bound of the delay derivative.

In [17], we proved extensions of the celebrate Razumikhin's theorem for a general family of time-varying continuous and discrete-time nonlinear systems. Our results include a novel "strictification" technique for converting a nonstrict Lyapunov function into a strict one. We also provided new constructions of Lyapunov-Krasovskii functionals that can be used to prove robustness to perturbations. Our examples include a key model from identification theory, and they show how our method can sometimes allow broader classes of delays than the results in the literature.

7.8. Systems with Long Delays

Participants: Frederic Mazenc, Silviu Niculescu, Michael Malisoff [LSU,USA], Jerome Weston [LSU,USA], Ali Zemouche, Bin Zhou [Harbin Institute of Technology], Qingsong Liu [Harbin Institute of Technology].

We solved several problems of observer and control designs pertaining to the fundamental (and difficult) case where a delay in the input is too long for being neglected.

In [35], we studied the stabilization of linear systems with both state and input delays where the input delay can be arbitrarily large but exactly known. Observer-predictor based controllers are designed to predict the future states so that the input delay can be properly compensated. Necessary and sufficient conditions guaranteeing the stability of the closed-loop system are provided in terms of the stability of some simple linear time-delay systems refereed to as observer-error systems. Moreover, linear matrix inequalities are used to design both the state feedback gains and observer gains. Finally, a numerical example illustrates that the proposed approaches are more effective and safe to implement than the existing methods.

In [57], for a particular family of systems, we constructed observers in the case where the measured variables are affected by the presence of a point-wise time-varying delay. The key feature of the proposed observers is that the size of their gains is proportional to the inverse of the largest value taken by the delay. The main result is first presented in the case of linear chain of integrators and next is extended to nonlinear systems with specific nonlinearities (systems of feedforward form).

Two of our works are devoted to the development of the prediction technique based on sequential predictors. Let us recall that one of the key advantages of this method is that it circumvents the problem of constructing and estimating distributed terms in the control laws: instead of using distributed terms, our approach to handling longer delays is to increase the number of predictors. In [61], we provided a significant generalization of our previous results to cases with arbitrarily large feedback delay bounds, and where, in addition, current values of the plant state are not available to use in the sequential predictors. We illustrate our work in a pendulum example. In [18], we provided a new sequential predictors approach for the exponential stabilization of linear time-varying systems. Our method allows arbitrarily large input delay bounds, pointwise time-varying input delays and uncertainties. We obtain explicit formulas to find lower bounds for the number of required predictors.

7.9. Nonlinear Observer Design via LMIs

Participants: Ali Zemouche, Rajesh Rajamani [University of Minneapolis, USA], Hieu Trinh [Deakin University, Australia], Yan Wang [University of Minneapolis, USA], Michel Zasadzinski [CRAN], Hugues Rafaralahy [CRAN], Boulaid Boulkroune [Flanders Make, Lommel, Belgium], Gridsada Phanomchoeng [Chulalongkorn University,Thailand], Khadidja Chaib-Draa [University of Luxembourg], Mohamed Darouach [CRAN], Marouane Alma [CRAN], Holger Voos [University of Luxembourg].

- Observer Design for Lipschitz and Monotonic nonlinear systems using LMIs:

  New LMI (Linear Matrix Inequality) design techniques have been developed to address the problem of designing performant observers for a class of nonlinear systems. The developed techniques apply to both locally Lipschitz as well as monotonic nonlinear systems, and allows for nonlinear functions in both the process dynamics and output equations [59], [34]. The LMI design conditions obtained are less conservative than all previous results proposed in literature for these classes of nonlinear systems. By judicious use of Young's relation, additional degrees of freedom are included in the observer design. These additional decision variables enable improvements in the feasibility of the obtained LMI. Several recent results in literature are shown to be particular cases of the more general observer design methodology developed in this paper. Illustrative examples are used to show the effectiveness of the proposed methodology. The application of the method to slip angle estimation in automotive applications is discussed and experimental results are presented. Although this application was the main motivation of this work, the proposed techniques have been applied to an anaerobic digestion model for different contexts [43], [44], [45].

- HG/LMI Observer:

  A new high-gain observer design method with lower gain compared to the standard high-gain observer was proposed [62]. This new observer, called "HG/LMI" observer is obtained by combining the standard high-gain methodology with the LPVLMI-based technique. Through analytical developments, it is shown how the new observer provides a lower gain. A numerical example was used to illustrate the performance of the new "HG/LMI" observer that we can call "smart high-gain observer". The aim of this research is the application of this new observer design to estimate some variables in vehicle applications and other real-world applications.
- Dual Spatially Separated Sensors for Enhanced Estimation:

  Inspired by the function of spatially separated sensory organs found in nature, we explored the use of dual spatially separated sensors for enhanced estimation in modern engineering applications [26]. To illustrate the interest of dual spatially separated sensors, some real applications have been considered: 1) Adaptive parameter and state estimation in magnetic sensors; 2) Estimation of an unknown disturbance input in an automotive suspension; 3) Separation of inputs based on their direction of action in a digital stethoscope. Both analytical observer design developments and experimental evaluation of the results have been provided.

7.10. Observer-Based Stabilization of Uncertain Nonlinear Systems

Participants: Ali Zemouche, Rajesh Rajamani [University of Minneapolis, USA], Yan Wang [University of Minneapolis, USA], Fazia Bedouhne [University of Tizi-Ouzou, Algeria], Hamza Bibi [University of Tizi-Ouzou, Algeria], Abdel Aitouche [CRIStAL, Lille].

- Relaxed LMI conditions for switched systems and LPV systems:

  By exploiting the Finsler's lemma in a non-standard way, we derived new LMI conditions. This technique has been applied to linear switched systems with uncertain parameters [10], [40] and LPV (Linear Parameter Varying) systems with inexact parameters [39], respectively. In each case, the Finsler's inequality is exploited in a convenient way to get additional decision variables which render the LMIs less conservative than those existing in the literature. In addition to analytical comparisons, several numerical examples have been used to show the superiority of the proposed new LMI conditions.
- From LMI relaxations to sequential LMI algorithm:

  Recently, motivated by autonomous vehicle control problem, a robust observer based estimated state feedback control design method for an uncertain dynamical system that can be represented as a LTI system connected with an IQC-type nonlinear uncertainty was developed [28]. Different from existing design methodologies in which a convex semidefinite constraint is obtained at the cost of conservatism and unrealistic assumptions, the design of the robust observer state feedback controller is formulated in this paper as a feasibility problem of a bilinear matrix inequality (BMI) constraint. Unfortunately, the search for a feasible solution of a BMI constraint is a NP hard problem in general. The applicability of the linearization method, such as variable change method or congruence transformation, depends on the specific structure of the problem at hand and cannot be generalized. A new sequential LMI optimization method to search for a feasible solution was established. In the application part, a vehicle lateral control problem is presented to demonstrate the applicability of the proposed algorithm to a real-world estimated state feedback control design.

7.11. Analysis of PWA control of discrete-time linear dynamics in the presence of variable input delay

Participants: Sorin Olaru, Mohammed Laraba [CentraleSupélec], Silviu Niculescu.

We have addressed the robustness of a specific class of control laws, namely the piecewise affine (PWA) controllers, defined over a bounded region of the state-space. More precisely, we were interested in closed-loop systems emerging from linear dynamical systems controlled via feedback channels in the presence of varying transmission delays by a PWA controller defined over a polyhedral partition of the state-space. We exploit the fact that the variable delays are inducing some particular model uncertainty. Our objective was to characterize the delay invariance margins: the collection of all possible values of the time-varying delays for which the positive invariance of the corresponding region is guaranteed with respect to the closed-loop dynamics. These developments are proving to be useful for the analysis of different design methodologies and, in particular, for model predictive control (MPC) approaches. The proposed delay margin describes the admissible transmission delays for an MPC implementation. From a different perspective, the delay margin further characterizes the fragility of an embedded MPC implementation via the on-line optimization and subject to variable computational time.

7.12. On the precision in polyhedral partition representation and the fragility of PWA control

Participants: Sorin Olaru, Rajesh Koduri [CentraleSupélec], Pedro Rodriguez [CentraleSupélec].

Explicit model predictive control (EMPC) solves a multi-parametric Quadratic Programming (mp-QP) problem for a class of discrete-time linear system with linear inequality constraints. The solution of the EMPC problem in general is a piecewise affine control function defined over non-overlapping convex polyhedral regions composing a polyhedral partition of the feasible region. In this work, we considered the problem of perturbations on the representation of the vertices of the polyhedral partition. Such perturbations may affect some of the structural characteristics of the PWA controller such as *non-overlapping within the regions* or *the closed-loop invariance*. We first showed how a perturbation affects the polyhedral regions and evoked the overlapping within the modified polyhedral regions. The major contribution of this work is to analyze to what extend the non-overlapping and the invariance characteristics of the PWA controller can be preserved when the perturbation takes place on the vertex representation. We determined a set called sensitivity margin to characterize for admissible perturbation preserving the non-overlapping and the invariance property of the controller. Finally, we show how to perturb multiple vertices sequentially and reconfigure the entire polyhedral partition

7.13. Convex Lifting: Theory and Control Applications

Participants: Sorin Olaru, Martin Gulan [STU, Bratislava, Slovaquie], Ngoc Anh Nguyen [J. Kepler Univ., Linz, Austria], Pedro Rodriguez [CentraleSupélec].

We introduced the *convex lifting* concept which was proven to enable significant implementation benefits for the class of piecewise affine controllers. Accordingly, two different algorithms to construct a convex lifting for a given polyhedral/polytopic partition were presented. These two algorithms rely on either the vertex or the halfspace representation of the related polyhedra. Also, we introduced an algorithm to refine a polyhedral partition, which does not admit a convex lifting, into a convexly liftable one. Furthermore, two different schemes are put forward to considerably reduce both the memory footprint and the runtime complexity which play a key role in implementation of piecewise affine controllers. These results have been illustrated via a numerical example and a complexity analysis.

7.14. Attitude control

Participants: Frederic Mazenc, Maruthi Akella [Univ. of Texas,USA], Sungpil Yang [Univ. of Texas,USA].

In [31], we addressed adaptive control of specific Euler-Lagrange systems: rigid-body attitude control, and the n-link robot manipulator. For each problem, the model parameters are unknown but the lower bound of the smallest eigenvalue of the inertia matrix is assumed to be known. The dynamic scaling Immersion and Invariance (I&I) adaptive controller is proposed to stabilize the system without employing a filter for the regressor matrix. A scalar scaling factor is instead implemented to overcome the integrability obstacle that arises in I&I adaptive control design. First, a filter-free controller is proposed for the attitude problem such that the rate feedback gain is proportional to the square of the scaling factor in the tracking error dynamics. The gain is then shown to be bounded through state feedback while achieving stabilization of the tracking error. The dynamic scaling factor increases monotonically by design and may end up at a finite but arbitrarily large value. However, by introducing three more dynamic equations, the non-decreasing scaling factor can be removed from the closed-loop system. Moreover, the behavior of dynamic gain is dictated by design parameters so that its upper bound is limited by a known quantity and its final value approaches the initial value. A similar approach for the dynamic gain design is also applied to a filter-dependent controller where a filter for the angular rate is utilized to build a parameter estimator. Unlike the filter-free design, the filter-dependent controller admits a constant gain for the rate feedback while the dynamic scaling factor rather appears in the filter. Finally, the proposed design is applied to robot manipulator systems. Spacecraft attitude and 2-link planar robot tracking problems are considered to demonstrate the performance of the controllers through simulations.

The work [32] builds on the preliminary results by generalizing to the tracking case and some further analysis of the filter-free case. Extending the strictification technique, a partially strict Lyapunov function is constructed toward establishment of stability and ultimate boundedness properties for the closed-loop system. With known upper bounds of the magnitude of measurement errors, disturbance torques, and parameter uncertainties, a feasible range for the feedback gains is derived in terms of bounds on the initial conditions in such a way to ensure asymptotic convergence of all closed-loop signals to within a residual set. In spite of the nonlinear structure of the kinematics and dynamics of the problem, however, the closed-loop system is rigorously analyzed through the standard Lyapunov analysis methods. This is achieved owing to the fact that the strictified Lyapunov function allows us to deal with this nontrivial problem in a standard way. As the passivity-based controller is not new for the attitude control problem, the key contribution of this paper is a theoretical analysis of the ideal case design in the presence of uncertainties through Lyapunov stability analysis.

7.15. Active Vibration Control of thin structures

Participants: Islam Boussaada, Silviu-Iulian Niculescu, Sami Tliba [L2S], Hakki Unal [Anadolu University], Toma Vyhlidal [Czech Technical University].

The problem of active vibrations damping of thin mechanical structures is a topic that has received great attention by the control community for several years, especially, when actuators and sensors are based on piezoelectric materials. For mechanical structures that are deformable, piezoelectric materials are used as strain sensors or strain actuators. With an appropriate controller, they allow to achieve shape control or the active damping of multi-modal vibrations thanks to their very large bandwidth. In this area, the major challenge is the design of controllers able to damp the most vibrating modes in a specified low-frequency bandwidth while ensuring robustness against high-frequency modes, outside the bandwidth of interest, often unmodelled or weakly modelled. The inherent feature of this kind of systems is that they arise robustness issues when they are tackled with finite dimensional control tools. A delayed state-feedback control strategies based on rightmost spectral values assignment allowing a fast vibration damping are proposed in [69], [41], [11].

7.16. Automatic Train Supervision for a CBTC Suburban Railway Line Using Multiobjective Optimization

Participants: Guillaume Sandou, Juliette Pochet [SNCF], Sylvain Baro [SNCF].

Communication-based train control (CBTC) systems have been deployed on subway lines to increase capacities on existing infrastructures. For the same purpose, CBTC systems are to be deployed on suburban railway lines where operating principles and constraints are significantly different. A regulation method for CBTC trains on a suburban line has been developed. This method is designed to combine CBTC functionalities with suburban operating principles. It includes a traffic management method in station, and a rescheduling method in case of disturbances. The proposed regulation method is integrated into the railway system simulation tool SIMONE developed by SNCF. This simulation tool includes models of the whole CBTC system, as well as the classic signaling system, train dynamics and railway infrastructures. Models of these different agents are described. The integration of the proposed regulation method into the tool SIMONE allows evaluating performances while taking into account the functional complexity of a CBTC railway system. The approach is illustrated with a realistic case: simulations of a CBTC traffic on the urban part of a railway line in the Paris region network are described. The proposed regulation method shows interesting results in disturbed situations according to the railway operating principles [60].

7.17. A Distributed Consensus Control Under Disturbances for Wind Farm Power Maximization

Participants: Guillaume Sandou, Nicolo Gionfra [CentraleSupélec], Houria Siguerdidjane [CentraleSupélec], Damien Faille [EDF], Philippe Loevenbruck [CentraleSupélec].

We have addressed the problem of power sharing among the wind turbines (WTs) belonging to a wind farm. The objective is to maximize the power extraction under the wake effect, and in the presence of wind disturbances. Because of the latter, WTs may fail in respecting the optimal power sharing gains. These are restored by employing a consensus control among the WTs. In particular, under the assumption of discrete-time communication among the WTs, we propose a distributed PID-like consensus approach that enhances the rejection of the wind disturbances by providing the power references to the local WT controllers. The latter are designed by employing a novel feedback linearization control that, acting simultaneously on the WT rotor speed and the pitch angle, guarantees the tracking of general deloaded power references. The obtained results are validated on a 6-WT wind farm example. [50].

7.18. Distributed Particle Swarm Optimization Algorithm for the Optimal Power Flow Problem

Participants: Guillaume Sandou, Nicolo Gionfra [CentraleSupélec], Houria Siguerdidjane [CentraleSupélec], Damien Faille [EDF], Philippe Loevenbruck [CentraleSupélec].

The distributed optimal power flow problem has been addressed. No assumptions on the problem cost function, and network topology are needed to solve the optimization problem. A particle swarm optimization algorithm is proposed, based on Deb's rule to handle hard constraints. Moreover, the approach enables to treat a class of distributed optimization problems, via a population based algorithm, in which the agents share a common optimization variable. A simulation example is provided, based on a 5-bus electric grid. [51].

7.19. Chemostat

Participants: Frederic Mazenc, Michael Malisoff [LSU,USA], Gonzalo Robledo [Univ. de Chile,Chile].

A chemostat is a fundamental bioreactor used to study the behavior of microorganisms. Many different types of chemostats exist, and many different types of models represent them.

We studied in [56] a chemostat model with an arbitrary number of competing species, one substrate, and constant dilution rates. We allowed delays in the growth rates and additive uncertainties. Using constant inputs of certain species, we derived bounds on the sizes of the delays that ensure asymptotic stability of an equilibrium when the uncertainties are zero, which can allow persistence of multiple species. Under delays and uncertainties, we provided bounds on the delays and on the uncertainties that ensure a robustness property of input-to-state stability with respect to uncertainties.

In [16], we provided a new control design for chemostats, under constant substrate input concentrations, using piecewise constant delayed measurements of the substrate concentration. Our growth functions can be uncertain and are not necessarily monotone. The dilution rate is the control. We use a new Lyapunov approach to derive conditions on the largest sampling interval and on the delay length to ensure asymptotic stabilization properties of a componentwise positive equilibrium point.

7.20. Qualitative/quantitative analysis of a delayed chemical model

Participants: Islam Boussaada, Silviu-Iulian Niculescu, Hakki Unal [Anadolu University].

The Belousov-Zhabotinsky reaction is a complex chemical reaction exhibiting sustained oscillations observed in some real biological oscillators. However, its oscillatory behavior is represented by a simple mechanism, called the Oregonator. A qualitative/quantitative analysis of a two-delay Oregonator based chemical oscillator is considered where the delay effect in dynamics is investigated; the existence of positive equilibrium point, the stability and boundedness of solutions for positive initial conditions are explored [27].

7.21. Mathematical Modelling of Acute Myeloid Leukemia

Participants: Catherine Bonnet, Jean Clairambault [MAMBA project-team], François Delhommeau [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Walid Djema, Emilia Fridman [Tel-Aviv University], Pierre Hirsch [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Frédéric Mazenc, Hitay Özbay [Bilkent University].

The ALMA3-project is about the modeling and analysis of healthy and unhealthy cell population dynamics, with a particular focus on hematopoiesis, which is the process of blood cell production and continuous replenishment. We point out that medical research is now looking for new combined targeted therapies able to overcome the challenge of cancer cells (e.g. to stop overproliferation, to restore normal apoptosis rates and differentiation of immature cells, and to avoid the high toxicity effects that characterize heavy non-selective chemotherapy). In that quest, the ultimate goal behind mathematical studies is to provide some inputs that should help biologists to suggest and test new treatment, and to contribute within multi-disciplinary groups in the opening of new perspectives against cancer. Thus, our research project is imbued within a similar spirit and fits the expectations of a better understanding of the behavior of healthy and unhealthy blood cell dynamics. It involve intensive collaboration with hematologists from Saint Antoine hospital in Paris, and aims to analyze the cell fate evolution in treated or untreated leukemia, allowing for the suggestion of new anti-leukemic combined chemotherapy.

In 2017, we have discussed some of the issues that are related to the modeling of the cell cycle, with particular insight into hematopoietic systems. For instance, i) we introduced and studied for the first time the effect of cell plasticity (dedifferentiation and transdifferentiation mechanisms) in the class of models that we focus on, and ii) we considered the effect of cell-arrest (i.e. some cells can be arrested during their cell-cycle) in models with several maturity stages. Stability features of the resulting biological models are highlighted, since systems trajectories reflect the most prominent healthy or unhealthy behaviors of the biological process under study. We indeed perform stability analysis of systems describing healthy and unhealthy situations, particularly in the case of acute myeloblastic leukemia (AML). More precisely, these are nonlinear time-delay systems that involve finite or infinite distributed delay terms, with possibly time-varying parameters. We pursue the objectives of earlier works in order to understand the interactions between the various parameters and functions involved in the mechanisms we study. Sometimes, we extend the stability analysis and the application of some already existing models, whereas news models and variants are other times introduced to cover novel biological evidences, such as: mutations accumulation and cohabitation between ordinary and mutated cells in niches, control and eradication of cancer stem cells, cancer dormancy and cell plasticity. In fact, the challenging problem that we are facing is to steadily extend both modelling and analysis aspects to constantly better represent this complex physiological mechanism, which is not yet fully understood. So, this year, we have progressed on our project and we have extended our works in order to develop the modeling and analysis aspects in cancer dormancy by including the effects of immuno-therapies in AML [48]. Lyapunov-like techniques have been used in this work in order to derive global or local exponential stability conditions for that class of differential-difference systems. Finally, in [49], we have modeled the role played by growth factors -these are hormone-like molecules- or drugs on the regulation of various biological features that are involved in hematopoiesis.

7.22. Analysis of Dengue Fever SIR Model with time-varying parameters

Participants: Stefanella Boatto [Univ Feder Rio de Janeiro], Catherine Bonnet, Frédéric Mazenc, Le Ha Vy Nguyen.

Dengue fever is an infectious viral disease occurring in humans that is prevalent in parts of Central and South America, Africa, India and South-east Asia and which causes 390 millions of infections worldwide. We continued this year our study on modeling of dengue epidemics.

We have first considered a SIR model with birth and death terms and time-varying infectivity parameter $\beta(t)$. In the particular case of a sinusoidal parameter, we showed that the average Basic Reproduction Number R_o, introduced in [Bacaër & Guernaoui, 2006], is not the only relevant parameter and we emphasized the rôle played by the initial phase, the amplitude and the period. For a (general) periodic infectivity parameter $\beta(t)$ a periodic orbit exists, as already proved in [Katriel, 2014]. In the case of a slowly varying $\beta(t)$ an approximation of such a solution is given, which is shown to be asymptotically stable under an extra assumption on the slowness of $\beta(t)$. For a non necessarily periodic $\beta(t)$, all the trajectories of the system are proved to be attracted into a tubular region around a suitable curve, which is then an approximation of the underlying attractor. Numerical simulations are given [68].

In other to study the effects of urban human mobility on Dengue epidemics, we have considered a SIR-network model (still with birth and death rates). The same model without these rates was introduced in [72].

In the case of constant infection rates, we first examine networks of two nodes. For arbitrary network topologies, some general properties of the equilibrium points are obtained. Then for several specific topologies, we derive explicit expressions of multiple equilibrium points and characterize their stability properties. We extend the study to networks with an arbitrary number of nodes and obtain sufficient conditions for global asymptotic stability of the disease-free equilibrium point.

In the case of time-varying infection rates and networks of arbitrary number of nodes, we introduce a specific topology which leads to a simplification of the network: the dynamics of the total population is described by the classical SIR model. This fact, together with the results of the team on the SIR model, allows a complete characterization of the stability properties of the system, especially the approximation of the epidemic attractor.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

A collaboration with SNCF on the supervision and rescheduling of a mixed CBTC traffic on a suburban railway line is currently undergoing (CIFRE).

A collaboration with EDF on the control of renewable energy parks is undergoing (financial support of a PhD student).

A collaboration with CEA and ADEME on the modelling and control of district heating networks is undergoing (financial support of a PhD student).

9. Partnerships and Cooperations

9.1. Regional Initiatives

DIGITEO Project (DIM LSC) ALMA3

Project title: Mathematical Analysis of Acute Myeloid Leukemia (AML) and its treatments

September 2014 - August 2017

Coordinator: Catherine Bonnet

Other partners: Inria Paris-Rocquencourt, France, L2S, France, UPMC, St Antoine Hospital Paris

Abstract: this project follows the regional projects ALMA (2010-2014) and ALMA2 (2011-2013). Starting from the work of J. L. Avila Alonso's PhD thesis in ALMA the aim of this project was to provide a refined coupled model of healthy and cancer cell dynamics in AML whose (stability) analysis may enable evaluation of polychemiotherapies delivered in the case of AML which have a high level of Flt-3 duplication (Flt-3-ITD).

9.2. National Initiatives

9.2.1. Industrial-Academic Institute

Guillaume Sandou is the head of the RISEGrid Institute. The Institute is dedicated to the study, modelling and simulation of smart electric distribution grids and their interactions with the whole electric power system. It is located in CentraleSupélec and gathers about 20 people (academic and industrial researchers, PhD students, post-doctoral researchers).

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

Program: **ITN**

Project acronym: TEMPO

Project title: Training in Embedded Predictive Control and Optimization

Duration: January 2014 - January 2018

Coordinator: Tor Arne Johanson; with Sorin Olaru (as French PI)

Other partners: U. Frieburg, Oxford, Imperial College; NTNU Trondheim; STUBA Bratislava; EPFL Lausanne; KU Leuven, Renault, ABB, Ampyx Power

Abstract: TEMPO is an international PhD program for highly motivated young scientists, where state-of-the-art research is combined with a comprehensive training program. The network is funded by the European Communityâs Seventh Framework program. The European Commission wants to make research careers more attractive to young people and therefore offers early-stage researchers (ESRs) a PhD program the opportunity to improve their research skills, join established research teams and enhance their career prospects via the Marie Curie Initial Training Networks (ITN) in the area of Embedded Predictive Control and Optimization.

9.3.2. Collaborations in European Programs, Except FP7 & H2020

Program: **PHC BOSPHORE 2016 (Turkey)**

Project title: Robust Control of Time Delayed Linear Parameter Varying Systems via Switched Controllers.

Duration: January 2016 - December 2017

Coordinator: Frédéric Mazenc (France), Hitay Özbay (Turkey).

Abstract: The main goal of this project is to develop computational algorithms for robust controller design for different classes of time delay systems appearing in various engineering applications such as chemical processes, transportation systems and communications networks. The participants will consider control problems of significant practical implications in this area: (i) developing new computational techniques for simple (low order) reliable and scalable decentralized controllers for control of (and control over) networks; and (ii) reducing conservatism in recently developed dwell-time based stability results for the analysis of switched time delay systems. Moreover, design of scalable low order controllers for reducing the effect of time delays is an important problem investigated in this project. One of the objectives of this collaboration is to generalize the design techniques already developed by the French and Turkish teams to larger classes of time delay systems, in particular multi-input-multi-output (MIMO) systems with time varying delays.

Program: **PHC BRANCUSI 2017 (Romania)**

Project acronym: ProCo

Project title: Systems with propagation: New approaches in control design for oscillation quenching

Duration: January 2016 - December 2018

Coordinator: Islam Boussaada (France) et Daniela Danciu (Romania)

Abstract: The project aims to building a unitary framework for the modeling, the analysis and the control of distributed-parameters systems (DPS) described by hyperbolic partial differential equations in one space variable and non-standard boundary conditions. This main objectives are modeling of DPS and the corresponding functional differential equations, the construction of reduced-order models approximating DPS by both numerical and computational modeling, the design of new control methods for oscillations quenching in DPS.

Program: **PHC CARLSO FINLEY 2017 (Cuba)**

Project title: MODELISATION ET COMMANDE POUR LE PROCESSUS DE CRYOCONSERVATION.

Duration: June 2017 - December 2017

Coordinator: Sorin Olaru (France), Marcos Martinez Montero (Turkey).

Abstract: The aim of this project is to initiat a collaboration on subjects related to the mathematical modelling of the dynamics involved in the cryopreservations process. In particular, the viability analysis of the vegetal material subject to cryogeny is one of the main objectives. The approach will realy on the evaluation electric leakage properties.

Program: **COST Action**

Project acronym: FRACTAL

Project title: Fractional-order systems; analysis, synthesis and their importance for future design

Duration: November 2016 - October 2020

Coordinator: Jaroslav Koton Czech Republic

Abstract: Fractional-order systems have lately been attracting significant attention and gaining more acceptance as generalization to classical integer-order systems. Mathematical basics of fractional-order calculus were laid nearly 300 years ago and since that it has gained deeply rooted mathematical concepts. Today, it is known that many real dynamic systems cannot be described by a system of simple differential equation or of integer-order system. In practice we can encounter such systems in electronics, signal processing, thermodynamics, biology, medicine, control theory, etc. The Action will favor scientific advancement in above mentioned areas by coordinating activities of academic research groups towards an efficient deployment of fractal theory to industry applications.

9.4. International Initiatives

Catherine Bonnet is the co-supervisor together with André Fioravanti of a PhD student of Unicamp (Brazil).

Frédéric Mazenc is the co-supervisor together with Hitay Ozbay of a PhD Student of Bilkent University (Turkey).

9.4.1. Inria International Partners

9.4.1.1. Informal International Partners

- College of Mathematics and Information Science, Shaanxi Normal University, China
- School of Control Science and Engineering, Dalian University of Technology, Dalian, China
- Louisiana State University, Baton Rouge, USA
- School of Electrical Engineering at the Tel-Aviv University, Israel

- The University of Texas at Austin, Dept. of Aerospace Engineering & Engineering Mechanics, USA
- Blikent University, Turkey
- Universidad de Chile, Chile
- School of Mathematics, University of Leeds, U.K.
- University Federale Rio de Janeiro, Brazil
- UNICAMP, Brazil
- Kyoto University, Japan

9.4.2. Participation in Other International Programs

9.4.2.1. International Initiatives

STADE

Title: Stability and Dichotomies in Differential Equations (Ordinary & Delay).

International Partners (Institution - Laboratory - Researcher):

Universidad de Chile (Chile) - Mathematics Department - Gonzalo Robledo

Universidad de la Republica Uruguay (Uruguay) - Faculty of Engineering - Pablo Monzon

Duration: 2016 - 2017

Start year: 2016

See also: http://www.stade.cl/pages/list.html

The ship-flags of this project are the concepts of dichotomy and stability in an ODE & DDE framework. We intend to study some theoretical and applied problems involving these concepts and its relations. In particular, converse stability results (expressed in the existence of density functions), feedback stabilization, stability in delay differential equations and some applications to bioprocesses.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

Stefanella Boatto, Federale University Rio de Janeiro, Brazil, 1 January-31 December.

André Fioravanti, UNICAMP, Sao Paulo, Brazil, 7 January-28 February.

Yutaka Yamamoto, Kyoto University, Japan, 17 May -2 August.

Hitay Ozbay, Bilkent University, Turkey, 15 November 2017 - 18 November 2017.

9.5.2. Visits to International Teams

Stefanella Boatto visited the Department of Mathematics, Universidade de Lisboa, Portugal, 19-23 June 2017. Frédéric Mazenc visited the Department of Mathematics of the Louisiana State University, Baton Rouge USA, 2 April - 14 April 2017, the Departamento de Ingeniaria de Control y Robotico of the Universidad Nacional Autonoma de Mexico, Mexico-city 14 August 2017 - 16 August and 18 August - 27 August 2017, the Laboratoire Franco-Mexicain d'Informatique et d'Automatique (LAFMIA), Mexico-City, 17 August 2017, Universidad de Chile, Santiago de Chile, 15 October 2017 to 28 October 2017.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. Member of the Organizing Committees

- Frederic Mazenc has organized with Pablo Monzon, Alvaro Castaneda and Gonzalo Robledo the "Workshop and Spring School on Stability and Dichotomies on Differential and Delay Equations" of the Universidad de Chile, Facultad de Ciencias, October 17-26, 2017.

10.1.2. Scientific Events Selection

- Frederic Mazenc and Ali Zemouche were Associate Editor for the conferences 2018 American Control Conference, Milwaukee, USA, and the 56th IEEE Conference on Decision and Control, Melbourne, USA, (2017).
- Frederic Mazenc was Associate Editor for the European Control Conference, Limassol, Cyprus (2018).
- Ali Zemouche has co-organized three invited sessions in international conferences (*IEEE−ACC 2018*, *Workshop on Advanced Control and Diagnosis, ACD 2017*, *International Conference on Systems and Control, ICSC 2017*).

10.1.2.1. Chair of Conference Program Committees

- Ali Zemouche was an invited session chair of the "14th International Workshop on Advanced Control and Diagnosis, ACD 2017", which was held at Politehnica University of Bucharest, Romania, from 16 to 17 November 2017.
- International Society of Difference Equations (ISDE) Board of Directors has favorably voted for the proposal submitted by Sorin Olaru to organize the International Conference on Difference Equations and Applications (ICDEA) in 2021 in Paris Saclay.
- Guillaume Sandou is a member of the Program Committee of the IEEE Symposium on Computational Intelligence in Production and Logistics Systems, 2017, Hawai, USA

10.1.2.2. Member of the Conference Program Committees

- Catherine Bonnet was a member of the *Comité International Scientifique* de MADEV17, Rabat, Marocco.
- Frederic Mazenc is member (Associate Editor) of the *Control Editorial Board - IEEE - CSS*.
- Ali Zemouche is member (Associate Editor) of the *Control Editorial Board - IEEE - CSS*.
- Ali Zemouche was involved in the Technical Program Committee and International Program Committee of the following international conferences:
  - IEEE American Control Conference, ACC 2017;
  - Workshop on Advanced Control and Diagnosis, ACD 2017;
  - International Conference on Systems and Control, ICSC 2017;
  - Australian and New Zealand Control Conference, ANZCC 2017;
  - International Conference on Electrical Engineering and Control Applications, ICEECA 2017.

10.1.2.3. Reviewer

The team reviewed papers for several international conferences including IEEE Conference on Decision and Control, IEEE American Control Conference, European Control Conference, IFAC World Congress.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

Frederic Mazenc is member of the editorial boards (Associate Editor) of the following journals:

- IEEE Transactions on Automatic Control;
- European Journal of Control;
- Journal of Control and Decision.

Sorin Olaru is member of the editorial boards (Associate Editor) of the following journals:

- IMA Journal of Mathematical Control and Information;
- IEEE CSS-Letters.

Ali Zemouche is member of the editorial boards (Associate Editor) of the following journals:

- SIAM Journal on Control and Optimization;
- European Journal of Control;
- Cogent Engineering.
- Managing Guest Editor for a **Special Issue** in European Journal of Control
  - *Title:* Advanced Control and Observers for Complex Systems via LMIs
  - *Organizers:* *Ali Zemouche et al.*
  - *url:* *https://www.journals.elsevier.com/european-journal-of-control/call-for-papers/special-issue-on-advanced-control-and-observer-design-for-no*

10.1.3.2. Reviewer - Reviewing Activities

The team reviewed papers for several journals including SIAM Journal on Control and Optimization, Automatica, IEEE Transactions on Automatic Control, IEEE Control Systems Magazine, Systems and Control Letters.

10.1.4. Invited Talks

Stefanella Boatto gave a talk entitled *'Modeling epidemics dynamics due to Aedes mosquitoes : the example of Rio de Janeiro and how to approximate an epidemic attractor*, Université de Bordeaux, 22 Dec 2017, a talk entitled *'The N-body pronlem on surfaces, Maxwell laws and the axioms of Mechanics'*, Fluid mechanics seminar, Dept. Mechanical Engineering , Universitat Rovira i Virgili, Tarragona, Spain, 30 June 2017, a talk entitled *'SIR-Network model : epidemics dynamics in a city & climate variations'*, Seminar of Analysis and Differential Equations, Dept. of Mathematics, University of Lisbon, Portugal, 20 June 2018, a talk entitled *' The N-body pronlem on surfaces, Maxwell laws and the axioms of Mechanics'*, IA Seminar, Physics Dept., Universirsity of Lisbon, Portugal 22 June 2017, Vortex Dynamics Group, a talk entitled *"N-body on surfaces of revolution: the rôle played by curvature and topology '*, School of Mathematics and Statistics, University of St. Andrews, UK, February 2017.

Catherine Bonnet and Frédéric Mazenc gave a talk entitled *'Modeling and Analysis of Cell Dynamics in Acute Myeloid Leukemia*, Institute of Disease Modeling, Seattle, USA, 30 May 2017.

Ali Zemouche gave a talk entitled *'Nonlinear observer design for Lipschitz systems'*, University of Toulon (IUT de Toulon, France).

Ali Zemouche gave a talk entitled *'Observer-based stabilization of uncertain nonlinear systems via LMIs'*, Deakin University (Geelong, Australia).

Frédéric Mazenc gave a talk entitled *Model reduction and predictor control*, the Departamento de Ingeniaria de Control y Robotico of the Universidad Nacional Autonoma de Mexico, Mexico-city, August 2017.

Frédéric Mazenc gave several talks and lectures to the "Workshop and Spring School on Stability and Dichotomies on Differential and Delay Equations" of the Universidad de Chile, Facultad de Ciencias, Santiago de Chile, October 17-26, 2017.

Frédéric Mazenc was one of the speakers of the tutorial session entitled "Tutorial on time-delay and sampled-data systems" organized by Alexandre Seuret and Emilia Fridman in the IFAC World Congress of Toulouse, 9-14 July 2017. The title of his talk was *Model reduction and predictor control*.

10.1.5. Leadership within the Scientific Community

Catherine Bonnet is a member of the IFAC Technical Committees *Distributed Parameter Systems* and *Biological and Medical Systems*. She is a member of the SIAG/CST (SIAM Activity group Control System Theory) steering committee (2015-2017) and a member of the management committee of the COST Action FRACTAL (2016-2020).

Sorin Olaru is a member of the IFAC Technical Committees *Robust Control* and the IFAC CSS TC on *hybrid systems*.

Ali Zemouche is member of the IFAC Technical Committee *Non-Linear Control Systems*.

10.1.6. Scientific Expertise

Catherine Bonnet is a member of the Evaluation Committee of Inria since September 2015.

Since 2014, Frédéric Mazenc is an expert for the FNRS (Belgium). His mission consists in evaluating research projects funded by this institution.

Since 2012, Frédéric Mazenc is a, expert for the ANVUR (National Agency for the Evaluation of Universities and Research Institutes, Italy). His mission consists in evaluating the contribution of Italian scientists.

Since 2011, Frédéric Mazenc is a, expert for the Romanian National Council for Development and Innovation (Romania). His mission consists in evaluating research projects funded by the this institution.

10.1.7. Research Administration

Catherine Bonnet is a member of the administration council of the association *Femmes et Mathématiques*, of the Parity Committee of Inria and of the *Cellule veille et prospective* of Inria (both created in 2015).

In 2017, Frédéric Mazenc was president of the commission scientifique du CRI Saclay-Ile-de-France. In 2017, Frédéric Mazenc was member of the Bureau du Comité des Projets du CRI Saclay-Ile-de-France. Since October 2017, he is Correspondant Inria Saclay A.M.I.E.S., http://www.agence-maths-entreprises.fr/

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master : Stefanella Boatto, Challenges in Biomathematical Modelling, 3h, M1, CentraleSupélec

Licence : Walid Djema, Computer Architecture and Assembly Programming, 40h, L1, University Paris-Saclay

Licence : Walid Djema, Computer Sciences project, 24h, L1, University Paris-Saclay

Master : Dina Irofti, Industrial IT services, Java, Networks, 64h, M1 and M2 University Paris-Sud

Master : Dina Irofti, Control Theory, Mathematics and Numericam Analysis, M1, 54h, ESIEE

Doctorat : Frederic Mazenc, introduction to the ordinary differential equations, Lypunov design, control and observation of nonlinear dynamical systems, 21h, PhD, International Graduate School on Control of the EECI, CentraleSupelec

Doctorat : Frederic Mazenc, Stability and Dichotomies on Differential and Delay Equations, 3 h, PhD, Universidad de Chile, Facultad de Ciencias

Licence : Sorin Olaru, Automatic Control, 8h , M1, SUPELEC, France

Licence : Sorin Olaru, Signals and systems, 8h , L3, SUPELEC, France

Licence : Sorin Olaru, Embedded systems, 8h , M1, Centrale Paris, France

Licence : Sorin Olaru, Numerical methods and Optimization, 24h, niveau M1, SUPELEC, France

Licence : Sorin Olaru, Hybrid systems, 16h, M2, SUPELEC, France

Licence : Guillaume Sandou, Signals and Systems, 87h, L3, CentraleSupélec

Licence : Guillaume Sandou, Mathematics and programming, 18h, L3, CentraleSupélec

Master : Guillaume Sandou, Automatic Control, 8h, M1, CentraleSupélec

Master : Guillaume Sandou, Numerical methods and optimization, 28h, M1 and M2, CentraleSupélec

Master : Guillaume Sandou, Modelling and system stability analysis, 21h, M2, CentraleSupélec

Master : Guillaume Sandou, Control of energy systems, 22h, M2, CentraleSupélec

Master : Guillaume Sandou, Robust control and mu-analysis, 9h, M2, CentraleSupélec

Master : Guillaume Sandou, Systems identification, 32h, M2, ENSTA

Master : Guillaume Sandou, System Analysis, 22h, M2, Ecole des Mines de Nantes

DUT : Ali Zemouche, Java Programing, 24 HTD, 2^{éme} année DUT (Bac + 2), University of Lorraine, France

10.2.2. Supervision

PhD in progress : Saeed Ahmed, Bilkent University, Stability analysis and control of switched systems with time-delay. Supervisor : Hitay Ozbay. Co-supervisor : Frédéric Mazenc.

PhD in progress : Nadine Aoun, Modélisation de réseaux de chaleur et gestion avancée multi-échelles de la production, de la distribution et de la demande. Modeling and multi-scale advanced management of production, distribution and demand in district heating networks. Supervisor: Guillaume Sandou.

PhD in progress : Caetano Cardeliquio, Stability and stabilization of (possibly fractional) systems with delays. French Supervisor : Catherine Bonnet, Brazilian Supervisor : André Fioravanti.

PhD : Walid Djema, Understanding Cell Dynamics in Cancer from Control and Mathematical biology Standpoints - Particular Insights in the Modeling and Analysis Aspects in Hematopoietic Systems and Leukemia, Université Paris-Saclay, 21 November 2017. Supervisor : Catherine Bonnet. Co-supervisors : Jean Clairambault and Frédéric Mazenc.

PhD : Dina Irofti, Delay effects: a journey from multi-agent systems to genetic networks, Université Paris-Saclay, 18 July 2017. Supervisor : Silviu Niculescu. Co-supervisor : Islam Boussaada.

PhD in progress : Mohamed Lotfi Derouiche, Sur l'optimisation par métaheuristiques avancées de lois de commande prédictive non linéaire. On the optimization of nonlinear predictive control laws using advanced metaheuristics algorithms. Supervisor: Soufienne Bouallegue, Joseph Haggége et Guillaume Sandou.

PhD in progress : Nicolo Gionfra, Optimisation du pilotage d'un parc d'énergies renouvelables avec stockage et du réseau de distribution sous-jacent. Optimization of the control of a park of renewable sources considering strorage means and distribution network. Supervisors: Houria Siguerdidjane et Guillaume Sandou.

PhD in progress : Juliette Pochet, Analyse de performance et de résilience d'une ligne de type RER équipée d'un automatisme CBTC. Analysis of the performance of a RER line with CBTC trains. Supervisor: Guillaume Sandou.

PhD in progress : Jean Mercat, Modele predictif des objets d'une scene routiere ; application à la sélection robuste des cibles pour les ADAS. Supervisor: Guillaume Sandou.

PhD in progress : Maxime Pouilly-Cathelain, Commande adaptative temps réel vis-a-vis de critères multiples de haut niveau. Supervisor : Guillaume Sandou.

10.2.3. Juries

Catherine Bonnet was a member of several recruiting committees: Junior Researcher competition in Inria Grenoble - Rhône-Alpes and Bordeaux - Sud-Ouest and Professor competition at Université de Nancy.

Catherine Bonnet was the President of the PhD Defense juries of Yacine Boukal *'Observation et commande des systèmes dynamiques d'ordre non entier'*, 16 October 2017, Université de Nancy and of Jin Chi *'Stability analysis of systems with delay-dependant coefficients'*, L2S, CentraleSupelec, 21 November 2017.

Sorin Olaru was a reviewer for the PhD thesis of JULIAN BARREIRO GOMEZ at Universiy of Catalunya in Barcelona.

Sorin Olaru was a reviewer of the HDR thesis of Christophe Louembet.

Guillaume Sandou was a reviewer of Khaleb Laib PhD, Analyse hierarchisee de a robustesse des systemes incertains de grade dimension.

Guillaume Sandou was a reviewer of Damien Casetta PhD, Modele d'aide à la conduite de réseaux de froid.

Frederic Mazenc was a reviewer of the PhD thesis of Mohamed Maghenem, *'Commande en formation de véhicules autonomes'* 05 July 2017, L2S, Centralesupelec.

Frederic Mazenc was a reviewer of the PhD thesis of Luis Borja Rosales, *'Stabilization of a class of nonlinear systems with passivity properties'*, 06 July 2017, L2S, Centralesupelec.

Frederic Mazenc was a reviewer of the PhD thesis of first half of the Phd ("suivi mi-parcours") of Mohamed Kahelras, *'Observation Problem for Different Classes of Nonlinear Delayed Systems'*, 9 October 2017, L2S, Centralesupelec.

10.3. Popularization

Catherine Bonnet gave a talk in the *Promenades Mathématiques* series of the event *Femmes en maths : une équation lumineuse* for female high school students, IHP, Paris, 19 december 2017. She met several groups of female high school students at the event *Femmes en maths : une équation lumineuse* IHP, Paris, january 2017. She met several groups of high school students at the event *Forum des métiers* in Lycée Hoche, Versailles, February 2017.

11. Bibliography

Major publications by the team in recent years

[1] C. Bonnet, A. R. Fioravanti, J. R. Partington.*Stability of Neutral Systems with Commensurate Delays and Poles Asymptotic to the Imaginary Axis*, in "SIAM Journal on Control and Optimization", March 2011, vol. 49, n^{o} 2, p. 498-516, https://hal.inria.fr/hal-00782325.

[2] C. Bonnet, J. Partington.*Stabilization of some fractional delay systems of neutral type*, in "Automatica", 2007, vol. 43, p. 2047–2053.

[3] M. Malisoff, F. Mazenc.*Constructions of Strict Lyapunov Functions*, Communications and Control Engineering Series, Springer-Verlag London Ltd., 2009.

[4] F. Mazenc, M. Malisoff, S.-I. Niculescu.*Reduction Model Approach for Linear Time-Varying Systems with Delays*, in "IEEE Transactions on Automatic Control", 2014, vol. 59, n^{o} 8, p. 2068–2014.

[5] W. MICHIELS, S.-I. NICULESCU.*Stability and Stabilization of Time-Delay Systems. An Eigenvalue-Based Approach*, Advances in Design and Control, SIAM: Philadelphia, 2007, vol. 12.

[6] S.-I. NICULESCU.*Delay Effects on Stability: a Robust Control Approach*, Lecture Notes in Control and Information Sciences, Springer, 2001, vol. 269.

[7] S. OLARU, D. DUMUR.*Avoiding constraints redundancy in predictive control optimization routines*, in "IEEE Trans. Automat. Control", 2005, vol. 50, n^{o} 9, p. 1459–1465.

[8] G. SANDOU.*Particle swarm optimization: an efficient tool for the design of automatic control law*, in "European Control Conference", Budapest, Hungary, August 23rd-26th 2009.

Publications of the year

Articles in International Peer-Reviewed Journal

[9] F. ADIB YAGHMAIE, R. SU, F. L. LEWIS, S. OLARU.*Bipartite and cooperative output synchronizations of linear heterogeneous agents: A unified framework*, in "Automatica", March 2017, vol. 80, p. 172 - 176 [*DOI :* 10.1016/J.AUTOMATICA.2017.02.033], https://hal-centralesupelec.archives-ouvertes.fr/hal-01509304.

[10] H. BIBI, F. BEDOUHENE, A. ZEMOUCHE, H.-R. KARIMI, H. KHELOUFI.*Output feedback stabilization of switching discrete-time linear systems with parameter uncertainties*, in "Journal of The Franklin Institute", September 2017, vol. 354, n^{o} 14, p. 5895-5918 [*DOI :* 10.1016/J.JFRANKLIN.2017.07.027], https://hal.archives-ouvertes.fr/hal-01567361.

[11] I. BOUSSAADA, S.-I. NICULESCU, S. TLIBA, T. VYHLÍDAL.*On the Coalescence of Spectral Values and its Effect on the Stability of Time-delay Systems: Application to Active Vibration Control*, in "Procedia IUTAM", 2017, vol. 22, p. 75 - 82 [*DOI :* 10.1016/J.PIUTAM.2017.08.011], https://hal-centralesupelec.archives-ouvertes.fr/hal-01656723.

[12] I. BOUSSAADA, S. TLIBA, S.-I. NICULESCU, H. U. ÜNAL, T. VYHLÍDAL.*Further remarks on the effect of multiple spectral values on the dynamics of time-delay systems. Application to the control of a mechanical system*, in "Linear Algebra and its Applications", 2018, p. 1-16 [*DOI :* 10.1016/J.LAA.2017.11.022], https://hal-centralesupelec.archives-ouvertes.fr/hal-01657659.

[13] J. CHEN, P. FU, C. F. MÉNDEZ BARRIOS, S.-I. NICULESCU, H. ZHANG.*Stability Analysis of Polynomially Dependent Systems by Eigenvalue Perturbation*, in "IEEE Transactions on Automatic Control", 2017, vol. 62, 8, accepté le 23 décembre 2016 [*DOI :* 10.1109/TAC.2016.2645758], https://hal-centralesupelec.archives-ouvertes.fr/hal-01431810.

[14] W. DJEMA, F. MAZENC, C. BONNET.*Stability analysis and robustness results for a nonlinear system with distributed delays describing hematopoiesis*, in "Systems and Control Letters", 2017, vol. 102, p. 93 - 101 [*DOI :* 10.1016/J.SYSCONLE.2017.01.007], https://hal.inria.fr/hal-01627125.

[15] X.-G. LI, S.-I. NICULESCU, A. CELA, L. ZHANG, X. LI.*A frequency-sweeping framework for stability analysis of time-delay systems*, in "IEEE Transactions on Automatic Control", 2017, vol. 64, 16, accepté le 21 octobre 2016 [*DOI :* 10.1109/TAC.2016.2633533], https://hal-centralesupelec.archives-ouvertes.fr/hal-01431725.

[16] F. MAZENC, J. HARMAND, M. MALISOFF.*Stabilization in a Chemostat with Sampled and Delayed Measurements and Uncertain Growth Functions*, in "Automatica", April 2017, vol. 78, p. 241-249 [*DOI :* 10.1016/J.AUTOMATICA.2016.12.035], https://hal.archives-ouvertes.fr/hal-01604883.

[17] F. MAZENC, M. MALISOFF.*Extensions of Razumikhin's Theorem and Lyapunov-Krasovskii Functional Constructions for Time-Varying Systems with Delay* *, in "Automatica", April 2017, vol. 78, p. 1-13 [*DOI :* 10.1016/J.AUTOMATICA.2016.12.005], https://hal.inria.fr/hal-01660097.

[18] F. MAZENC, M. MALISOFF.*Stabilization and Robustness Analysis for Time-Varying Systems with Time-Varying Delays using a Sequential Subpredictors Approach* *, in "Automatica", August 2017, vol. 82, p. 118-127 [*DOI :* 10.1016/J.AUTOMATICA.2017.04.020], https://hal.inria.fr/hal-01660103.

[19] F. MAZENC, M. MALISOFF.*Stabilization of Nonlinear Time-Varying Systems through a New Prediction Based Approach*, in "IEEE Transactions on Automatic Control", June 2017, vol. 62, n^{o} 6, 8 [*DOI :* 10.1109/TAC.2016.2600500], https://hal.inria.fr/hal-01389859.

[20] F. MAZENC, M. MALISOFF, S.-I. NICULESCU.*Stability and Control Design for Time-Varying Systems with Time-Varying Delays using a Trajectory-Based Approach*, in "SIAM Journal on Control and Optimization", January 2017, vol. 55, n^{o} 1, p. 533 - 556 [*DOI :* 10.1137/15M1027838], https://hal.inria.fr/hal-01660100.

[21] F. MAZENC, M. MALISOFF, H. OZBAY.*Stability and robustness analysis for switched systems with time-varying delays*, in "SIAM Journal on Control and Optimization", 2017, p. 1-24, https://hal.inria.fr/hal-01660106.

[22] W. MICHIELS, I. BOUSSAADA, S.-I. NICULESCU.*An explicit formula for the splitting of multiple eigenvalues for nonlinear eigenvalue problems and connections with the linearization for the delayeigenvalue problem*, in "SIAM Journal on Matrix Analysis and Applications", June 2017, vol. 38, n^{o} 2, p. 599–620 [*DOI :* 10.1137/16M107774X], https://hal-centralesupelec.archives-ouvertes.fr/hal-01558169.

[23] L. H. V. NGUYEN, C. BONNET.*Stabilization of MISO fractional systems with delays*, in "Automatica", 2017, vol. 83, p. 337 - 344 [*DOI :* 10.1016/J.AUTOMATICA.2017.06.032], https://hal.inria.fr/hal-01627128.

[24] L. H. V. NGUYEN.*A unified approach for the H_∞-stability analysis of classical and fractional neutral systems with commensurate delays*, in "SIAM Journal on Control and Optimization", 2018, https://hal.archives-ouvertes.fr/hal-01679070.

[25] N. A. NGUYEN, S. OLARU, P. RODRIGUEZ-AYERBE, M. KVASNICA.*Convex liftings-based robust control design*, in "Automatica", March 2017, vol. 77, p. 206 - 213 [*DOI :* 10.1016/J.AUTOMATICA.2016.11.031], https://hal-centralesupelec.archives-ouvertes.fr/hal-01509316.

[26] R. RAJAMANI, Y. WANG, G. D. NELSON, R. MADSON, A. ZEMOUCHE.*Observers with dual spatially separated sensors for enhanced estimation: Industrial, automotive, and biomedical applications*, in "IEEE Control Systems", June 2017, vol. 37, n^{o} 3, p. 42-58 [*DOI :* 10.1109/MCS.2017.2674439], https://hal.archives-ouvertes.fr/hal-01534700.

[27] H. U. UNAL, I. BOUSSAADA, S.-I. NICULESCU.*A Delay-Based Sustained Chemical Oscillator: Qualitative Analysis of Oregonator-Based Models*, in "IEEE Life Sciences Letters", September 2017, vol. 3, n^{o} 3, p. 9 - 12 [*DOI :* 10.1109/LLS.2017.2756834], https://hal-centralesupelec.archives-ouvertes.fr/hal-01656725.

[28] Y. WANG, R. RAJAMANI, A. ZEMOUCHE.*Sequential LMI approach for design of a BMI-based robust observer state feedback controller with nonlinear uncertainties*, in "International Journal of Robust and Nonlinear Control", 2018, p. 1-17 [*DOI :* 10.1002/RNC.3948], https://hal.archives-ouvertes.fr/hal-01575709.

[29] F. XU, S. OLARU, V. PUIG, C. OCAMPO-MARTINEZ, S.-I. NICULESCU.*Sensor-fault tolerance using robust MPC with set-based state estimation and active fault isolation*, in "International Journal of Robust and Nonlinear Control", May 2017, vol. 27, n^{o} 8, p. 1260 - 1283 [*DOI :* 10.1002/RNC.3627], https://hal-centralesupelec.archives-ouvertes.fr/hal-01512371.

[30] F. XU, V. PUIG, C. OCAMPO-MARTINEZ, S. OLARU, S.-I. NICULESCU.*Robust MPC for actuator-fault tolerance using set-based passive fault detection and active fault isolation*, in "International Journal of Applied Mathematics and Computer Science", March 2017, vol. 27, n^{o} 1, p. 43-61 [*DOI :* 10.1515/AMCS-2017-0004], https://hal-centralesupelec.archives-ouvertes.fr/hal-01495636.

[31] S. YANG, M. AKELLA, F. MAZENC.*Dynamically Scaled Immersion and Invariance Adaptive Control for Euler-Lagrange Mechanical Systems*, in "Journal of Guidance, Control, and Dynamics", November 2017, vol. 40, n^{o} 11, p. 2844-2856 [*DOI :* 10.2514/1.G002177], https://hal.inria.fr/hal-01662280.

[32] S. YANG, F. MAZENC, M. AKELLA.*Ultimate Boundedness Results for Noise-Corrupted Quaternion Output Feedback Attitude Tracking Controllers*, in "Journal of Guidance, Control, and Dynamics", December 2017, vol. 40, n^{o} 12, p. 3265 - 3273 [*DOI :* 10.2514/1.G002164], https://hal.inria.fr/hal-01660110.

[33] A. ZEMOUCHE, R. RAJAMANI, H. KHELOUFI, F. BEDOUHENE.*Robust observer-based stabilization of Lipschitz nonlinear uncertain systems via LMIs - Discussions and new design procedure*, in "International Journal of Robust and Nonlinear Control", July 2017, vol. 27, n^{o} 11, p. 1915-1939 [*DOI :* 10.1002/RNC.3644], https://hal.archives-ouvertes.fr/hal-01567359.

[34] A. ZEMOUCHE, R. RAJAMANI, G. PHANOMCHOENG, B. BOULKROUNE, H. RAFARALAHY, M. ZASADZINSKI.*Circle criterion-based $\mathcal{H}_\infty$ observer design for Lipschitz and monotonic nonlinear systems - Enhanced LMI conditions and constructive discussions*, in "Automatica", November 2017, vol. 85, p. 412-425 [*DOI :* 10.1016/J.AUTOMATICA.2017.07.067], https://hal.archives-ouvertes.fr/hal-01567360.

[35] B. ZHOU, Q. LIU, F. MAZENC.*Stabilization of Linear Systems with Both Input and State Delays by Observer-Predictors*, in "Automatica", September 2017, vol. 83, p. 368-377 [*DOI :* 10.1016/J.AUTOMATICA.2017.06.027], https://hal.inria.fr/hal-01660105.

Invited Conferences

[36] C. JIN, S.-I. NICULESCU, I. BOUSSAADA, K. GU.*Stability Analysis for Control Systems with Delay-Difference Approximation of Output Derivatives*, in "IFAC 2017 - 20th World Congress of the International Federation of Automatic Control", Toulouse, France, July 2017 [*DOI :* 10.1016/J.IFACOL.2017.08.1894], https://hal-centralesupelec.archives-ouvertes.fr/hal-01658188.

International Conferences with Proceedings

[37] C. BENNANI, F. BEDOUHENE, A. ZEMOUCHE, H. BIBI, A. AITOUCHE.*A modified two-steps LMI method to design observer-based controller for linear discrete-time systems with parameter uncertainties*, in "6th International Conference on Systems and Control, ICSC' 2017", Batna, Algeria, May 2017, https://hal.archives-ouvertes.fr/hal-01567355.

[38] H. BIBI, F. BEDOUHENE, A. ZEMOUCHE, A. AITOUCHE.*$\mathcal{H}_\infty$ observer-based stabilization of switched discrete-time linear systems*, in "6th International Conference on Systems and Control, ICSC 2017", Batna, Algeria, May 2017, https://hal.archives-ouvertes.fr/hal-01567356.

[39] H. BIBI, F. BEDOUHENE, A. ZEMOUCHE, H.-R. KARIMI.*A new LMI-based output feedback controller design method for discrete-time LPV systems with uncertain parameters*, in "20th IFAC World Congress, IFAC 2017", Toulouse, France, July 2017, https://hal.archives-ouvertes.fr/hal-01567352.

[40] H. BIBI, A. ZEMOUCHE, A. AITOUCHE, K. CHAIB DRAA, F. BEDOUHENE.*Robust observer-based H_∞ stabilization for a class of switched discrete-time linear systems with parameter uncertainties*, in "56th IEEE Conference on Decision and Control, CDC 2017", Melbourne, Australia, December 2017, https://hal.archives-ouvertes.fr/hal-01674803.

[41] I. BOUSSAADA, S. TLIBA, S.-I. NICULESCU.*A Delayed Feedback Controller for Active Vibration Control : A Rightmost-Characteristic Root Assignment based Approach*, in "ICSTCC 2017 - 21st International Conference on System Theory, Control and Computing", Sinaia, Romania, October 2017, https://hal-centralesupelec.archives-ouvertes.fr/hal-01658209.

[42] C. CARDELIQUIO, M. SOUZA, A. FIORAVANTI.*Stability Analysis and Output-Feedback Control Design for Time-Delay Systems ⋆*, in "IFAC World Congress", Toulouse, France, July 2017, https://hal.inria.fr/hal-01679094.

[43] K. CHAIB DRAA, M. ALMA, H. VOOS, A. ZEMOUCHE, M. DAROUACH.*LMI-Based invariant like nonlinear state observer for anaerobic digestion model*, in "25th Mediterranean Conference on Control and Automation, MED 2017", Valletta, Malta, July 2017, https://hal.archives-ouvertes.fr/hal-01567358.

[44] K. CHAIB DRAA, H. VOOS, M. ALMA, A. ZEMOUCHE, M. DAROUACH.*An LMI-based H_∞ discrete-time non linear state observer design for an anaerobic digestion model*, in "20th IFAC World Congress, IFAC 2017", Toulouse, France, July 2017, https://hal.archives-ouvertes.fr/hal-01567351.

[45] K. CHAIB DRAA, H. VOOS, M. ALMA, A. ZEMOUCHE, M. DAROUACH.*LMI-based $\mathcal{H}_\infty$ nonlinear state observer design for anaerobic digestion model*, in "25th Mediterranean Conference on Control and Automation, MED 2017", Valletta, Malta, July 2017, https://hal.archives-ouvertes.fr/hal-01567357.

[46] K. CHAIB DRAA, H. VOOS, M. ALMA, A. ZEMOUCHE, M. DAROUACH.*LMI-based discrete-time nonlinear state observer for an anaerobic digestion model*, in "6th International Conference on Systems and Control, ICSC 2017", Batna, Algeria, May 2017, https://hal.archives-ouvertes.fr/hal-01567354.

[47] K. CHAIB DRAA, H. VOOS, M. ALMA, A. ZEMOUCHE, M. DAROUACH.*Observer-based trajectory tracking for anaerobic digestion process*, in "14th International Workshop on Advanced Control and Diagnosis, ACD 2017", Bucarest, Romania, November 2017, p. 1-6, https://hal.archives-ouvertes.fr/hal-01683627.

[48] W. DJEMA, C. BONNET, J. CLAIRAMBAULT, F. MAZENC, P. HIRSCH, F. DELHOMMEAU.*Analysis of a Model of Dormancy in Cancer as a State of Coexistence Between Tumor and Healthy Stem Cells*, in "ACC 2017 - American Control Conference", Seattle, United States, IEEE, May 2017, p. 5135-5140 [*DOI :* 10.23919/ACC.2017.7963751], https://hal.inria.fr/hal-01677927.

[49] W. DJEMA, H. ÖZBAY, C. BONNET, E. FRIDMAN, F. MAZENC, J. CLAIRAMBAULT.*Analysis of Blood Cell Production under Growth Factors Switching*, in "IFAC 2017 - 20th World Congress of the International

Federation of Automatic Control", Toulouse, France, Elsevier, July 2017, vol. 50, n^{o} 1, p. 13312-13317 [*DOI :* 10.1016/J.IFACOL.2017.08.1331], https://hal.inria.fr/hal-01677914.

[50] N. GIONFRA, G. SANDOU, H. SIGUERDIDJANE, D. FAILLE, P. LOEVENBRUCK.*A Distributed Consensus Control Under Disturbances for Wind Farm Power Maximization*, in "56th IEEE Conference on Decision and Control", Melbourne, Australia, Proceedings of the 56th IEEE Conference on Decision and Control, December 2017, https://hal-centralesupelec.archives-ouvertes.fr/hal-01667884.

[51] N. GIONFRA, G. SANDOU, H. SIGUERDIDJANE, P. LOEVENBRUCK, D. FAILLE.*A Novel Distributed Particle Swarm Optimization Algorithm for the Optimal Power Flow Problem*, in "CCTA 2017 - 1st IEEE Conference on Control Technology and Applications", Kohala Coast, United States, Proceedings of the 1st IEEE Conference on Control Technology and Applications, August 2017, p. 1-8, https://hal-centralesupelec.archives-ouvertes.fr/hal-01667868.

[52] C. JIN, I. BOUSSAADA, S.-I. NICULESCU, K. GU.*An Overview of Stability Analysis of Systems with Delay Dependent Coefficients*, in "ICSTCC 2017 - 21st International Conference on System Theory, Control and Computing", Sinaia, Romania, October 2017, https://hal-centralesupelec.archives-ouvertes.fr/hal-01658201.

[53] F. MAZENC, S. AHMED, H. OZBAY.*State Feedback Stabilization of Switched Systems with Delay: Trajectory Based Approach **, in "ACC 2017 - The 2017 American Control Conference", Seattle, United States, IEEE, May 2017, p. 4540-4543 [*DOI :* 10.23919/ACC.2017.7963655], https://hal.inria.fr/hal-01660129.

[54] F. MAZENC, M. MALISOFF, L. BURLION, V. GIBERT.*Bounded Backstepping through a Dynamic Extension with Delay*, in "CDC 2017 - 56th IEEE Conference on Decision and Control", Melbourne, Australia, December 2017, p. 1-5, https://hal.inria.fr/hal-01660133.

[55] F. MAZENC, M. MALISOFF, H. OZBAY.*Stability Analysis of Switched Systems with Time-Varying Discontinuous Delays*, in "ACC 2017 - The 2017 American Control Conference", Seattle, United States, IEEE, May 2017, p. 1-5 [*DOI :* 10.23919/ACC.2017.7963758], https://hal.inria.fr/hal-01660128.

[56] F. MAZENC, M. MALISOFF, G. ROBLEDO.*Stability and Robustness Analysis for a Multi-Species Chemostat Model with Uncertainties*, in "ACC 2017 - American Control Conference", Seattle, United States, IEEE, May 2017, p. 1-5 [*DOI :* 10.23919/ACC.2017.7963267], https://hal.inria.fr/hal-01660127.

[57] F. MAZENC, A. ZEMOUCHE, S.-I. NICULESCU.*Observer with small gains in the presence of a long delay in the measurements*, in "56th IEEE Conference on Decision and Control, CDC 2017", Melbourne, Australia, December 2017, p. 1-6, https://hal.inria.fr/hal-01660130.

[58] M. T. NGUYEN, C. STOICA MANIU, S. OLARU.*Optimization-based control for Multi-Agent deployment via dynamic Voronoi partition*, in "20th IFAC World Congress", Toulouse, France, July 2017, p. 1864-1869, https://hal.archives-ouvertes.fr/hal-01562834.

[59] G. PHANOMCHOENG, A. ZEMOUCHE, R. RAJAMANI.*Real-time automotive slip angle estimation with extended H_∞ circle criterion observer for nonlinear output system*, in "American Control Conference, ACC 2017", Seattle, WA, United States, May 2017, https://hal.archives-ouvertes.fr/hal-01541110.

[60] J. POCHET, G. SANDOU, S. BARO.*Automatic Train Supervision for a CBTC Suburban Railway Line Using Multiobjective Optimization*, in "20th IEEE International Conference on Intelligent Transportation Systems",

Yokohama, Japan, Proceedings of the 20th IEEE International Conference on Intelligent Transportation Systems, October 2017, https://hal-centralesupelec.archives-ouvertes.fr/hal-01667907.

[61] J. WESTON, M. MALISOFF, F. MAZENC.*Sequential Predictors under Time-Varying Delays: Effects of Delayed State Observations in Dynamic Controller*, in "CDC 2017 - 56th IEEE Conference on Decision and Control", Melbourne, Australia, December 2017, p. 1-6, https://hal.inria.fr/hal-01660132.

[62] A. ZEMOUCHE.*Observer design for nonlinear systems by using high-gain and LPV/LMI-based technique*, in "American Control Conference, ACC 2017", Seattle, WA, United States, May 2017, https://hal.archives-ouvertes.fr/hal-01567353.

Conferences without Proceedings

[63] C. JIN, S.-I. NICULESCU, I. BOUSSAADA, K. GU.*Stability Analysis of Control Systems subject to Delay-Difference Feedback*, in "IFAC 2017 - 20th World Congress of the International Federation of Automatic Control", Toulouse, France, Elsevier, July 2017, vol. 50, n^{o} 1, p. 13330 - 13335 [*DOI :* 10.1016/J.IFACOL.2017.08.1894], https://hal-centralesupelec.archives-ouvertes.fr/hal-01656727.

[64] N. KACI, B. OULD BOUAMAMA, I. BOUSSAADA, A. DEBIANE.*Structural diagnosability analysis. Application to an induction motor*, in "SDEMPED 2017 - 11th IEEE International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives", Tinos, Greece, IEEE, August 2017, p. 1-7 [*DOI :* 10.1109/DEMPED.2017.8062374], https://hal-centralesupelec.archives-ouvertes.fr/hal-01658216.

[65] D. NIETO HERNANDEZ, J. A. ESCARENO, F. MENDES-BARRIOS, I. BOUSSAADA, D. LANGARICA CÓRDOBA.*Modeling and Control of an Interactive Tilt-rotor MAV*, in "RED-UAS 2017 - 4th International Workshop on Research, Education and Development on Unmanned Aerial Systems", Linkoping, Sweden, IEEE, October 2017, p. 270-275 [*DOI :* 10.1109/RED-UAS.2017.8101678], https://hal-centralesupelec.archives-ouvertes.fr/hal-01657896.

Scientific Books (or Scientific Book chapters)

[66] R. SIPAHI, S.-I. NICULESCU, F. M. ATAY.*Deterministic car-following traffic models. Delay effects and linear stability*, SpringerBriefs in Electrical and Computer Engineering (Control, Automation and Robotics), Springer : Heidelberg, 2017, accepté : novembre 2016, https://hal-centralesupelec.archives-ouvertes.fr/hal-01431320.

Other Publications

[67] S. BOATTO, C. BONNET, B. CAZELLES, F. MAZENC.*SIR model with time dependent infectivity parameter : approximating the epidemic attractor and the importance of the phase* , November 2017, Epidemics6– International Conference on Infectious Disease Dynamics, Poster, https://hal.inria.fr/hal-01689258.

[68] S. BOATTO, C. BONNET, B. CAZELLES, F. MAZENC.*SIR model with time dependent infectivity parameter : approximating the epidemic attractor and the importance of the initial phase*, January 2018, working paper or preprint, https://hal.inria.fr/hal-01677886.

References in notes

[69] I. BOUSSAADA, S. TLIBA, S.-I. NICULESCU, H. U. ÜNAL, T. VYHLÍDAL.*Further remarks on the effect of multiple spectral values on the dynamics of time-delay systems. Application to the control of a mechanical*

system, in "Linear Algebra and its Applications", 2018, p. 1-16 [*DOI :* 10.1016/J.LAA.2017.11.022], https://hal-centralesupelec.archives-ouvertes.fr/hal-01657659.

[70] H. KITANO.*Biological robustness*, in "Nature", 2004, vol. 5, p. 826–837.

[71] H. OZBAY, C. BONNET, A. FIORAVANTI.*PID Controller Design for Fractional-Order Systems with Time Delays*, in "Systems and Control Letters", January 2012, vol. 61, n^o 1, p. 18-23, http://hal.inria.fr/hal-00766062.

[72] L. M. STOLERMAN, D. COOMBS, S. BOATTO.*SIR-Network Model and Its Application to Dengue Fever*, in "SIAM Journal on Applied Mathematics", 2015, vol. 75, n^o 6, p. 2581–2609.

Project-Team EX-SITU

Extreme Situated Interacton

IN COLLABORATION WITH: Laboratoire de recherche en informatique (LRI)

IN PARTNERSHIP WITH:
Université Paris-Sud (Paris 11)

RESEARCH CENTER
Saclay - Île-de-France

THEME
Interaction and visualization

Table of contents

Project-Team EX-SITU

Creation of the Team: 2015 January 01, updated into Project-Team: 2017 July 01

Keywords:

Computer Science and Digital Science:

A5.1. - Human-Computer Interaction
A5.1.1. - Engineering of interactive systems
A5.1.2. - Evaluation of interactive systems
A5.1.5. - Body-based interfaces
A5.1.6. - Tangible interfaces
A5.1.7. - Multimodal interfaces

Other Research Topics and Application Domains:

B2.8. - Sports, performance, motor skills
B5.7. - 3D printing
B6.3.1. - Web
B6.3.4. - Social Networks
B9.2. - Art
B9.2.1. - Music, sound
B9.2.4. - Theater
B9.4. - Sciences

1. Personnel

Research Scientists

Wendy Mackay [Team leader, Inria, Principle Researcher, HDR]
Joanna Mcgrenere [Inria, Inria Chair, Jul 2017]
Theophanis Tsandilas [Inria, Researcher]

Faculty Members

Michel Beaudouin-Lafon [Univ. Paris-Sud, Professor, HDR]
Sarah Fdili Alaoui [Univ. Paris-Sud, Assistant Professor]
Cédric Fleury [Univ. Paris-Sud, Assistant Professor]

External Collaborators

Ghita Jalal [Univ. Paris-Sud, until Sep 2017]
Midas Nouwens [Aarhus University, from Feb 2017]
Can Liu [University College London, From Jan 2017]

Technical Staff

Romain Di Vozzo [Inria, from Dec 2017]
Olivier Gladin [Inria, until Nov 2017]
Nicolas Taffin [Inria, from Sep 2017]

PhD Students

Jessalyn Alvina [Inria]
Ignacio Avellino Martinez [Inria]
Marianela Ciolfi Felice [Univ. Paris-Sud]
Carla Griggio [Inria]

Shu Yuan Hsueh [Univ. Paris-Sud]
Germán Leiva [Inria]
Wanyu Liu [Telecom ParisTech]
Nolwenn Maudet [Univ. Paris-Sud]
Yujiro Okuya [Univ. Paris-Sud]
Jean-Philippe Riviere [Univ. Paris-Sud, from Oct 2017]
Philip Tchernavskij [Université Paris-Saclay]
Michael Wessely [Inria]

Post-Doctoral Fellows

John Maccallum [Inria]
Andrew Webb [Inria]

Administrative Assistant

Alexandra Merlin [Inria]

2. Overall Objectives

2.1. Overall Objectives

Interactive devices are everywhere: we wear them on our wrists and belts; we consult them from purses and pockets; we read them on the sofa and on the metro; we rely on them to control cars and appliances; and soon we will interact with them on living room walls and billboards in the city. Over the past 30 years, we have witnessed tremendous advances in both hardware and networking technology, which have revolutionized all aspects of our lives, not only business and industry, but also health, education and entertainment. Yet the ways in which we interact with these technologies remains mired in the 1980s. The graphical user interface (GUI), revolutionary at the time, has been pushed far past its limits. Originally designed to help secretaries perform administrative tasks in a work setting, the GUI is now applied to every kind of device, for every kind of setting. While this may make sense for novice users, it forces expert users to use frustratingly inefficient and idiosyncratic tools that are neither powerful nor incrementally learnable.

ExSitu explores the limits of interaction — how extreme users interact with technology in extreme situations. Rather than beginning with novice users and adding complexity, we begin with expert users who already face extreme interaction requirements. We are particularly interested in creative professionals, artists and designers who rewrite the rules as they create new works, and scientists who seek to understand complex phenomena through creative exploration of large quantities of data. Studying these advanced users today will not only help us to anticipate the routine tasks of tomorrow, but to advance our understanding of interaction itself. We seek to create effective human-computer partnerships, in which expert users control their interaction with technology. Our goal is to advance our understanding of interaction as a phenomenon, with a corresponding paradigm shift in how we design, implement and use interactive systems. We have already made significant progress through our work on instrumental interaction and co-adaptive systems, and we hope to extend these into a foundation for the design of all interactive technology — to create a *physics of interaction*.

3. Research Program

3.1. Research Program

We characterize Extreme Situated Interaction as follows:

Extreme users. We study extreme users who make extreme demands on current technology. We know that human beings take advantage of the laws of physics to find creative new uses for physical objects. However, this level of adaptability is severely limited when manipulating digital objects. Even so, we find that creative professionals—artistists, designers and scientists—often adapt interactive technology in novel and unexpected ways and find creative solutions. By studying these users, we hope to not only address the specific problems they face, but also to identify the underlying principles that will help us to reinvent virtual tools. We seek to shift the paradigm of interactive software, to establish the laws of interaction that significantly empower users and allow them to control their digital environment.

Extreme situations. We develop extreme environments that push the limits of today's technology. We take as given that future developments will solve “practical" problems such as cost, reliability and performance and concentrate our efforts on interaction in and with such environments. This has been a successful strategy in the past: Personal computers only became prevalent after the invention of the desktop graphical user interface. Smartphones and tablets only became commercially successful after Apple cracked the problem of a usable touch-based interface for the iPhone and the iPad. Although wearable technologies, such as watches and glasses, are finally beginning to take off, we do not believe that they will create the major disruptions already caused by personal computers, smartphones and tablets. Instead, we believe that future disruptive technologies will include fully interactive paper and large interactive displays.

Our extensive experience with the Digiscope WILD and WILDER platforms places us in a unique position to understand the principles of distributed interaction that extreme environments call for. We expect to integrate, at a fundamental level, the collaborative capabilities that such environments afford. Indeed almost all of our activities in both the digital and the physical world take place within a complex web of human relationships. Current systems only support, at best, passive sharing of information, e.g., through the distribution of independent copies. Our goal is to support active collaboration, in which multiple users are actively engaged in the lifecycle of digital artifacts.

Extreme design. We explore novel approaches to the design of interactive systems, with particular emphasis on extreme users in extreme environments. Our goal is to empower creative professionals, allowing them to act as both designers and developers throughout the design process. Extreme design affects every stage, from requirements definition, to early prototyping and design exploration, to implementation, to adaptation and appropriation by end users. We hope to push the limits of participatory design to actively support creativity at all stages of the design lifecycle. Extreme design does not stop with purely digital artifacts. The advent of digital fabrication tools and FabLabs has significantly lowered the cost of making physical objects interactive. Creative professionals now create hybrid interactive objects that can be tuned to the user's needs. Integrating the design of physical objects into the software design process raises new challenges, with new methods and skills to support this form of extreme prototyping.

Our overall approach is to identify a small number of specific projects, organized around four themes: *Creativity, Augmentation, Collaboration* and *Infrastructure*. Specific projects may address multiple themes, and different members of the group work together to advance these different topics.

4. Application Domains

4.1. Creative industries

We work closely with creative professionals in the arts and in design, including music composers, musicians, and sound engineers; painters and illustrators; dancers and choreographers; theater groups; graphic and industrial designers; and architects.

4.2. Scientific research

We work with creative professionals in the sciences and engineering, including neuroscientists and doctors; programmers and statisticians; chemists and astrophysicists; and researchers in fluid mechanics.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

- Wendy Mackay: Doctor Honoris Causa, Aarhus University (Denmark), September 2017
- Wanyu Liu: “1er Prix Doctorants ED STIC” of Université Paris-Saclay, November 2017, for “BIGNav: Information Theory meets Human-Computer Interaction”

BEST PAPERS AWARDS :

[20] **ACM CHI 2017 - International conference of Human-Computer Interaction**. W. LIU, R. LUCAS D'OLIVEIRA, M. BEAUDOUIN-LAFON, O. RIOUL.

6. New Software and Platforms

6.1. Platforms

6.1.1. WildOS

Participant: Michel Beaudouin-Lafon [correspondant].

WildOS is middleware designed to support applications that run in an interactive room, such as our WILD and WILDER rooms, with various interaction resources, including a tiled wall display, a motion tracking system, interactive tabletops, tablets, smartphones and custom-made or 3d printed interactive devices. The conceptual model of WildOS is a *platform*, such as the WILD or WILDER room, that can be described as a set of devices on which one or more applications can be run.

WildOS consists of a server running on a machine that has network access to all the machines involved in the platform, and a set of clients running on the various interaction resources, such as a display cluster or a tablet. Once *WildOS* is running, applications can be started and stopped and devices can be added to or removed from the platform.

WildOS relies on Web technologies, most notably Javascript and node.js, as well as node-webkit and HTML5. This makes it inherently portable (it is currently tested on Mac OS X and Linux). While applications can be developed only with these Web technologies, it is also possible to bridge to existing applications developed in other environments if they provide sufficient access for remote control. Sample applications include a web browser, an image viewer, a window manager, and the BrainTwister application developed in collaboration with neuroanatomists at NeuroSpin.

WildOS is used for several research projects at ExSitu and by other partners of the Digiscope project. It was also deployed on several of Google's interactive rooms in Mountain View, Dublin and Paris. It is available under on Open Source licence at https://bitbucket.org/mblinsitu/wildos.

- ACM: H.5.2 [User Interfaces]: Graphical user interfaces (GUI)
- Software benefit: helps development of multisurface applications.
- OS/Middleware: Crossplatform
- Required library or software: node.js, node-webkit
- Programming language: Javascript

6.1.2. Unity Cluster

Participants: Cédric Fleury [correspondant], Olivier Gladin.

Unity Cluster is middleware to distribute any Unity 3D (https://unity3d.com/) application on a cluster of computers that run in interactive rooms, such as our WILD and WILDER rooms, or immersive CAVES (Computer-Augmented Virtual Environments). Users can interact the the application with various interaction resources.

Unity Cluster provides an easy solution for running existing Unity 3D applications on any display that requires a rendering cluster with several computers. *Unity Cluster* is based on a master-slave architecture: The master computer runs the main application and the physical simulation as well as manages the input; the slave computers receive updates from the master and render small parts of the 3D scene. *Unity Cluster* manages data distribution and synchronization among the computers to obtain a consistent image on the entire wall-sized display surface.

Unity Cluster can also deform the displayed images according to the user's position in order to match the viewing frustum defined by the user's head and the four corners of the screens. This respects the motion parallax of the 3D scene, giving users a better sense of depth.

Unity Cluster is composed of a set of C Sharp scripts that manage the network connection, data distribution, and the deformation of the viewing frustum. In order to distribute an existing application on the rendering cluster, all scripts must be embedded into a Unity package that is included in an existing Unity project.

- ACM: C.2.4 [Distributed Systems]: Distributed applications, I.3.7 [3D Graphics and Realism]: Virtual reality
- Software benefit: adapts existing Unity 3D application to a rendering cluster of an interactive room.
- OS/Middleware: Crossplatform
- Required library or software: Unity 3D
- Programming language: C Sharp

6.1.3. WILDER

Participants: Michel Beaudouin-Lafon [correspondant], Cédric Fleury, Olivier Gladin.

WILDER (Figure 1) is our second experimental ultra-high-resolution interactive environment, which follows the WILD platform developed in 2009. It features a wall-sized display with seventy-five 20" LCD screens, i.e. a 5m50 x 1m80 (18' x 6') wall displaying 14 400 x 4 800 = 69 million pixels, powered by a 10-computer cluster and two front-end computers. The platform also features a camera-based motion tracking system that lets users interact with the wall, as well as the surrounding space, with various mobile devices. The display uses a multitouch frame (the largest of its kind in the world) to make the entire wall touch sensitive.

WILDER was inaugurated in June, 2015. It is one of the ten platforms of the Digiscope Equipment of Excellence and, in combination with WILD and the other Digiscope rooms, provides a unique experimental environment for collaborative interaction.

In addition to using WILD and WILDER for our research, we have also developed software architectures and toolkits, such as WildOS and Unity Cluster, that enable developers to run applications on these multi-device, cluster-based systems.

7. New Results

7.1. Fundamentals of Interaction

Participants: Michel Beaudouin-Lafon [correspondant], Marianela Ciolfi Felice, Sarah Fdili Alaoui, Cédric Fleury, Carla Griggio, Wanyu Liu, Wendy Mackay, Nolwenn Maudet, Philip Tchernavskij, Theophanis Tsandilas.

In order to better understand fundamental aspects of interaction, ExSitu studies interaction under extreme situations. We conduct in-depth observational studies and controlled experiments which contribute to theories and frameworks that unify our findings and help us generate new, advanced interaction techniques.

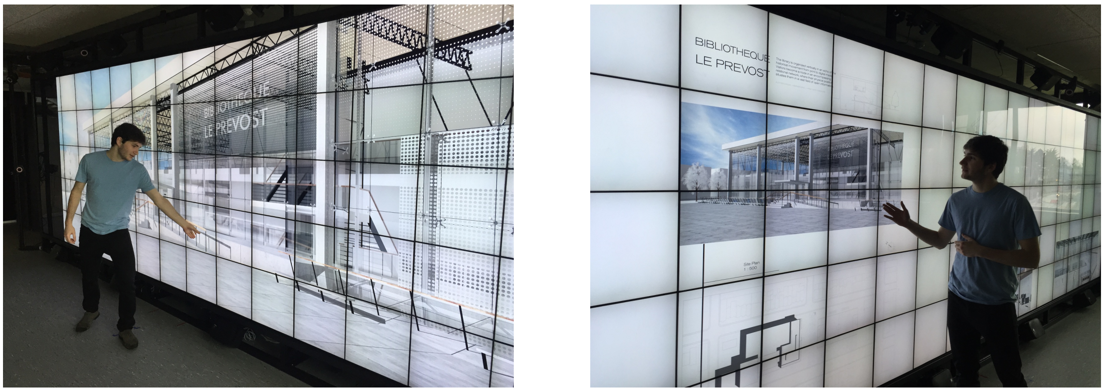


Figure 1. The WILDER platform.

On the theoretical side, in collaboration with Telecom ParisTech, we are bringing the tools and concepts from Information Theory to HCI. We conducted an information-theoretic analysis of human performance for command selection [21]. While a number of studies have focused on improving rapid command selection through novel interaction techniques, new interface design and innovative devices, user performance in this context has received little attention. We ran a controlled experiment to test the theory that the transmitted information from the user to the computer levels off as difficulty increases. Our reasoning is based on basic information-theoretic concepts such as entropy, mutual information and Fano's inequality. The important result is the bell-shaped behavior of the throughput as a function of command entropy, which shows that there is an optimal level of difficulty for a given input technique.

We also used the information-theoretic concept of mutual information, also known as information gain, in our BIG (Bayesian Information Gain) framework. We created *BIGnav* [20], a new multiscale navigation technique based on Bayesian Experimental Design where the criterion is to maximize the expected information gain from the next user input. In a controlled experiment, *BIGnav* was up to 40% faster than the standard pan-and-zoom technique. *BIGnav* creates a form of human-computer partnership (see below) where the computer challenges the user in order to maximize the amount of information extracted from the user's input. This work received a Best Paper Award at ACM CHI 2017, and the first prize for doctoral research from the Paris-Saclay doctoral school in computer science.

Finally, we continued our long-standing line of work on Fitts' law, with a novel analysis of minimal, as opposed to average, movement time in human aimed movement [18]. We showed that both metrics have a lot of support from theoretical and empirical perspectives and gave two examples, one in a controlled experiment and the other in a field study of pointing, where making the minimum versus average distinction is fruitful.

On the empirical side, we conducted two observational studies to better understand how people interact with technology. The first study [23] targeted expert graphic designers and their use of advanced computer tools. Traditional graphic design tools emphasize the grid for structuring layout. Interviews with professional graphic designers revealed that they use surprisingly sophisticated structures that go beyond the grid, which we call *graphical substrates*. These structures are not well supported by existing tools, so we developed two technology probes to explore how to embed graphical substrates into tools. *Contextify* lets designers tailor layouts according to each reader's intention and context, while *Linkify* lets designers create dynamic layouts based on relationships among content properties. We tested the probes with professional graphic designers, who all identified novel uses in their current projects. We incorporated their suggestions into *StyleBlocks*, a prototype that reifies CSS declarations into interactive graphical substrates. This work demonstrates that graphical substrates offer an untapped design space for tools that can help graphic designers generate personal layout structures.

The second study [30] targeted the operating system upgrade process that most users regularly have to go through to keep their system up to date. While current research has focused primarily on the security aspect of upgrades, we investigated the user's perspective of upgrading software. We found that users delay major upgrades by an average of 80 days, and an extensive field study revealed that very few participants prepare for upgrades (e.g., by backing up files), and over half had negative reactions to the upgrade process and other changes (e.g., bugs, lost settings, unwanted features). During the upgrade process, waiting times were too long, feedback was confusing or misleading, and few had clear mental models of what was happening. Moreover, users almost never mentioned security as a concern or reason for upgrading, while interviews with technical staff responsible for one organization's upgrades focused only on security and licensing, not user interface changes. This work shows that upgrades should be handled differently, offering users more control and decoupling security updates from the introduction of new features or the update of existing features.

These two sets of studies support our strong commitment to re-inventing interactive systems by identifying fundamental principles of interaction that unify, rather than separate, interaction styles in order to support the diversity of uses and users [33]. For example, most of our interactions with the digital world are mediated by apps: desktop, web, or mobile applications. Apps impose artificial limitations on collaboration among users, distribution across devices, and the changing procedures that constantly occur in real work. These limitations are partially due to the engineering principles of encapsulation and program-data separation, calling for new architectural principles [29]. Shareable dynamic media, which we have explored in our earlier work on Webstrates[5], provides an interesting approach as it blurs the limits between apps and documents and supports collaboration, distribution and flexibility as fundamental features [28]. In connection with these issues, we ran a workshop at the ACM CHI 2017 conference on HCI toolkits [37] where we discussed challenges and opportunities to develop new methods and approaches to design, evaluate, disseminate and share toolkits, as well as the technical, methodological and enabling role of toolkits for HCI research.

7.2. Human-Computer Partnerships

Participants: Wendy Mackay [correspondant], Jessalyn Alvina, Marianela Ciolfi Felice, Carla Griggio, Shu Yuan Hsueh, Wanyu Liu, John Maccallum, Nolwenn Maudet, Joanna Mcgrenere, Midas Nouwens, Andrew Webb.

ExSitu is interested in designing effective human-computer partnerships, in which expert users control their interaction with technology. Rather than treating the human users as the 'input' to a computer algorithm, we explore human-centered machine learning, where the goal is to use machine learning and other techniques to increase human capabilities. Much of human-computer interaction research focuses on measuring and improving productivity: our specific goal is to create what we call 'co-adaptive systems' that are discoverable, appropriable and expressive for the user. Jessalyn Alvina, under the supervision of Wendy Mackay, successfully defended her thesis, *Increasing The Expressive Power of Gesture-based Interaction on Mobile Devices* [38], on this topic.

We are interested in helping users create their own custom gesture-based commands for mobile devices. This raises two competing requirements: gestures must be both personally memorable for the user, while reliably recognizable by the system. We created two dynamic guides [22], *Fieldward* and *Pathward*, which use progressive feedforward to interactively visualize the “negative space” of unused gestures. The Pathward technique suggests four possible completions to the current gesture, whereas the Fieldward technique uses color gradients to reveal optimal directions for creating recognizable gestures (Figure 2). We ran a two-part experiment in which 27 participants each created 42 personal gesture shortcuts on a smartphone, using *Pathward*, *Fieldward* or *No Feedforward*. The Fieldward technique best supported the most common user strategy, i.e. to create a memorable gesture first and then adapt it to be recognized by the system. Users preferred the Fieldward technique to Pathward or No Feedforward, and remembered gestures more easily when using the technique. Dynamic guides can help developers design novel gesture vocabularies and support users as they design custom gestures for mobile applications.

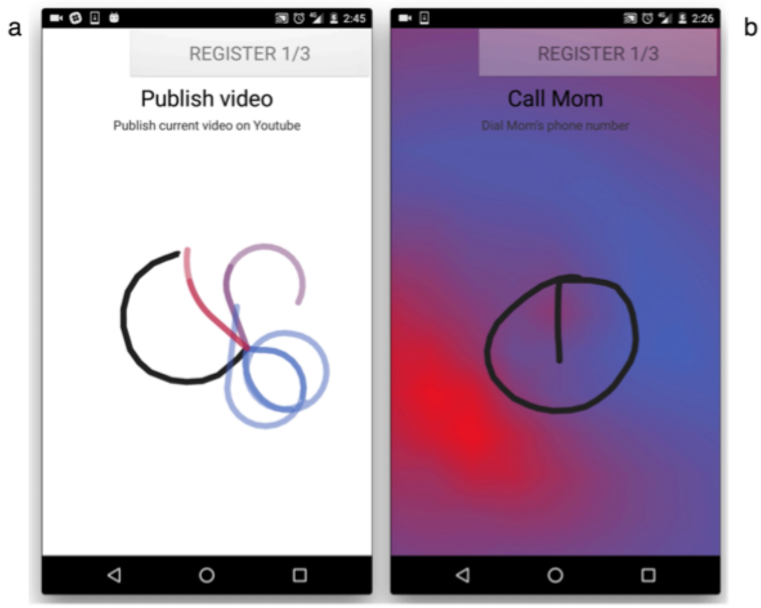


Figure 2. The Pathward (a) & Fieldward (b) dynamic guides help users create their own easy-to-remember gesture commands that are also recognizable by the system.

We are also interested in letting users use simple gestures to generate commands on a mobile device. *CommandBoard* [14] offers a simple, efficient and incrementally learnable technique for issuing gesture commands from a soft keyboard. We transform the area above the keyboard into a command-gesture input space that lets users draw unique command gestures or type command names followed by execute (Fig 3). Novices who pause see an in-context dynamic guide, whereas experts simply draw. Our studies show that *CommandBoard*'s inline gesture shortcuts are significantly faster (almost double) than markdown symbols and significantly preferred by users. We demonstrate additional techniques for more complex commands, and discuss trade-offs with respect to the user's knowledge and motor skills, as well as the size and structure of the command space. We filed a patent for the *CommandBoard* technique.

Figure 3. CommandBoard creates a new command gesture input space above a soft keyboard. Users can: a) type 'happy' and use a dynamic guide to style it as bold; b) type 'brightn', draw an execute gesture and adjust the brightness slider; c) type 'sans', choose 'sans mono' and draw an execute gesture to change the font; d) type 'color', select yellow in the marking menu to change the brush color.

In the context of an art-science project with the *n+1* theater group and the Théâtre de l'Agora d'Evry, we created an interactive installation that was exhibited at Fête de la Science, at the Agora d'Evry for an entire month, and at the Festival Curiositas. We were interested in understanding what makes public art installations interactive, so that they are engaging both for the individual user and the surrounding public. More specifically, we experimented with the principle of 'shaping' from behavioral psychology to create a human-computer partnership: an animated Santa character mirrors the exact movements of the user, but also offers different types of reinforcing or punishing feedback that in turn shapes the user's behavior (Figure 4). From the user's perspective, the user is always in control. Yet, from the system's perspective, the user moves through successive approximations to a specific desired behavior. Thus, we explore the dynamic nature of shared control between users and technology.

Finally, with *BIGnav* [20], we experimented with a different kind of partnership. *BIGnav* is a new multi-scale navigation technique based on Bayesian Experimental Design where the criterion is to maximize the information-theoretic concept of mutual information, also known as information gain. Rather than simply executing user navigation commands, *BIGnav* interprets user input to update its knowledge about the user's intended target. It then navigates to a new view that maximizes the information gain provided by the user's expected subsequent input. *BIGnav* creates a novel form of human-computer partnership, where the computer challenges the user in order to extract more information from the user's input, making interaction more efficient. We showed that *BIGnav* is significantly faster than conventional pan and zoom and requires fewer

Figure 4. The interactive Christmas window: the Santa character mimics the movements of the user in front of the window, but also uses reinforcement feedback to shape the movements of the user.

commands for distant targets, especially in non-uniform information spaces. We also applied *BIGnav* to a realistic application and showed that users can navigate to highly probable points of interest on a map with only a few steps.

7.3. Creativity

Participants: Sarah Fdili Alaoui [correspondant], Marianela Ciolfi Felice, Carla Griggio, Shu Yuan Hsueh, Ghita Jalal, Germán Leiva, John Maccallum, Wendy Mackay, Nolwenn Maudet, Joanna Mcgrenere, Midas Nouwens, Jean-Philippe Riviere, Nicolas Taffin, Philip Tchernavskij, Theophanis Tsandilas, Andrew Webb, Michael Wessely.

ExSitu is interested in understanding the work practices of creative professionals, particularly artists, designers, and scientists, who push the limits of interactive technology. Nolwenn Maudet, under the supervision of Wendy Mackay and Michel Beaudouin-Lafon, successfully defended her thesis, *Designing Design Tools* [40], on this topic. Her research includes observational studies of graphic designers and developers ([24] described in the *Collaboration* section below), as well as the creation of variety of creativity support tools to support professional designers [23] (described in the *Fundamentals of Interaction* section above).

We designed and evaluated computational models of movement's expressive qualities as defined in the framework of Laban Efforts [13] for dancer and movement practitioners. We included experts in Laban Movement Analysis (LMA) in our design process, in order to select a set of suitable multimodal sensors as well as to compute features that closely correlate to the definitions of Efforts in LMA. Evaluation of our model showed that multimodal data combining positional, dynamic and physiological information allows for a better characterization of Laban Efforts. Inspired by movement practices and dance, we designed an interactive sound installation that supports kinesthetic awareness of a participant's micro-movements [17] and discussed perspectives of such an installation from somatic practices and embodied cognition [16]. We discussed [25] the ethical and aesthetic implications of the appropriation of biomedical sensors in artistic practices, in particular dance. We also traced the history and new perspective of HCI in Dance and body based practices [11].

In collaboration with Inria Lille, we developed a versioning and annotation system for supporting collaborative, iterative design of mapping layers for digital musical instruments (DMIs) [31]. We also collaborated with Saarland University, TU Berlin and MIT to digitally fabricate Directional screens, devices and surfaces that maximize perceived image quality (e.g., resolution, brightness, and color reproduction) for large audiences [27]. Finally, Michael Wessely participated in the *MIT Summer School for Computational Fabrication and Smart Matter* and was then invited by its organizers to co-author an article [35] that presents and discusses the results of the summer school.

7.4. Collaboration

Participants: Cédric Fleury [correspondant], Ignacio Avellino Martinez, Michel Beaudouin-Lafon, Marianela Ciolfi Felice, Carla Griggio, Germán Leiva, Can Liu, Wendy Mackay, Nolwenn Maudet, Joanna Mcgrenere, Midas Nouwens, Yujiro Okuya.

ExSitu is interested in exploring new ways of supporting collaborative interaction, especially within and across large interactive spaces such as those of the Digiscope network (http://digiscope.fr/). Multi-touch wall-sized displays afford collaborative exploration of large datasets and re-organization of digital content. However, standard touch interactions, such as dragging to move content, do not scale well to large surfaces and were not designed to support collaboration, such as passing objects around. We created *CoReach* [19], a set of collaborative gestures that combine input from multiple users in order to manipulate content, facilitate data exchange and support communication. *Throw-and-catch* (Figure 5) lets users send digital objects to each other, *Preview* lets one user show content to another, and *SharedClipboard* lets users gather content. We conducted an observational study to inform the design of *CoReach*, and a controlled study showing that it reduced physical fatigue and facilitated collaboration when compared with traditional multi-touch gestures. A final study assessed the value of also allowing input through a handheld tablet to manipulate content from a distance.

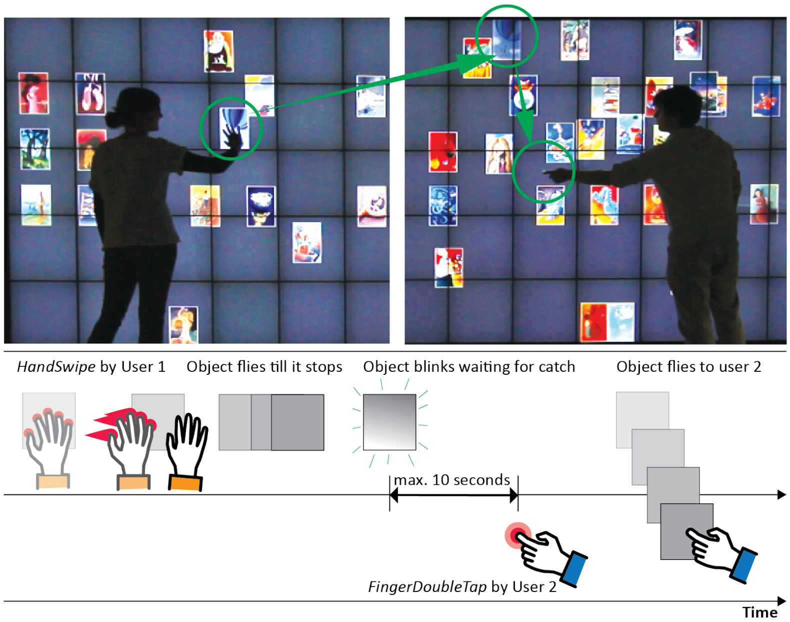


Figure 5. One of the three CoReach gestures: sending an object to the partner.

We also studied remote collaboration across wall-sized displays, where the challenge is to support audio-video communication among users as they move in front of the display. We created *CamRay* [15], a platform that uses camera arrays embedded in wall-sized displays (Figure 6) to capture video of users and present it on remote displays according to the users' positions. We investigated two settings: in *Follow-Remote*, the position of the video window follows the position of the remote user; in *Follow-Local*, the video window always appears in front of the local user. A controlled experiment showed that with *Follow-Remote*, participants were faster, used more deictic instructions, interpreted them more accurately, and used fewer words. However, some participants preferred the virtual face-to-face created by *Follow-Local* when checking for their partners' understanding. An ideal system should therefore combine both modes, in a way that does not hinder the collaborative process. Ignacio Avellino, under the supervision of Michel Beaudouin-Lafon and Cédric Fleury, successfully defended his thesis *Supporting Collaborative Practices Across Wall-Sized Displays with Video-Mediated Communication* [39] on this topic.

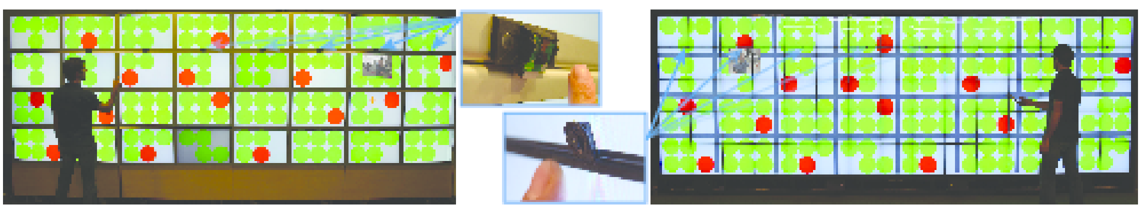


Figure 6. CamRay: the video cameras embedded in the wall-sized display (middle), and the two wall-sized displays (left, right) showing the video feed from the partner.

We are also interested in the collaboration between between professional interaction designers and software developers when they create novel interactive systems [24]. Although designers and developers have different skills and training, they need to collaborate closely to create interactive systems. Our studies highlighted the mismatches among their processes, tools and representations: We found that current practices create unnecessary rework and cause discrepancies between the original design and the implementation. We identified three types of design breakdowns: omitting critical details, ignoring edge cases, and disregarding technical limitations. In a follow-up study, we found that early involvement of the developer helped mitigate potential design breakdowns but new ones emerged as the project unfolded. Finally, we ran a participatory design session and found that the designer/developer pairs had difficulty representing and communicating pre-existing interactions. This work will inform our future work on tools for designers and developers of interactive systems, in the context of our overal theoretical framework.

We also studied the use of social networks, with an indepth study of how users communicate via multiple apps that offer almost identical functionality [26] We studied how and why users distribute their contacts within their app ecosystem. We found that the contacts in an app affect a user's conversations with other contacts, their communication patterns in the app, and the quality of their social relationships. Users appropriate the features and technical constraints of their apps to create idiosyncratic communication places, each with its own re- cursively defined membership rules, perceived purposes, and emotional connotations. Users also shift the boundaries of their communication places to accommodate changes in their contacts' behaviour, the dynamics of their relationships, and the restrictions of the technology. We argue that communi- cation apps should support creating multiple communication places within the same app, relocating conversations across apps, and accessing functionality from other apps.

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. MoveIT – Modeling the Speed/Accuracy Trade-Off of Human Aimed Movement with the Tools of Information Theory

Type: Ph.D. grant

Funding: DigiCosme Labex

Duration: 2015-2018

Coordinator: Olivier Rioul (Institut Mines Telecom)

Partners: Univ. Paris-Sud, Inria, CNRS, Institut Mines-Telecom

Inria contact: Michel Beaudouin-Lafon

Abstract: The goal of this project is to conduct fundamental studies of aimed movements based on information theory. The project studies the interaction phenomena involved in pointing, in order to discover novel, more effective pointing techniques. This project funds Wanyu Liu, a joint Ph.D. student between the COMELEC and VIA groups at Institut Mines Telecom and ExSitu.

8.1.2. SensoMotorCVE – Sensor-motor Interface for Collaborative Virtual Environments with Heterogeneous Devices: Application to Industrial Design

Type: Ph.D. grant

Funding: DigiCosme Labex

Duration: 2014-2017

Coordinator: Patrick Bourdot (LIMSI-CNRS)

Partners: Univ. Paris-Sud, Inria, CNRS

Inria contact: Cédric Fleury

Abstract: In the context of collaborative virtual environments, the goal of this project is to develop a sensorimotor interface model for CAD data manipulation that supports heterogeneous interactive systems such as wall-sized displays or immersive virtual reality rooms. This project funds Yujiro Okuya, a joint Ph.D. student between the VENISE group at LIMSI and ExSitu.

8.1.3. An Augmented-Reality System for Collaborative Physical Modeling and Design

Type: Equipment

Funding: STIC Paris-Saclay

Duration: 2017-2018

Coordinator: Theophanis Tsandilas

Partners: Univ. Paris-Sud, Inria

Inria contact: Theophanis Tsandilas

Abstract: The goal of the project is to develop an augmented-reality system to support collaboration over 3D models and enhance digital-fabrication approaches. It is a collaboration with the AVIZ group and provides funding (8k) for equipment.

8.1.4. Le Plateau des Recherches Infinies

Type: Equipment and subcontracting

Funding: Learning Center Paris-Saclay

Duration: 2017-2018

Coordinator: Michel Beaudouin-Lafon

Partners: Univ. Paris-Sud

Inria contact: Michel Beaudouin-Lafon

Abstract: The goal of this project (30k) is to create an interactive installation presenting the portraits of a hundred researchers from Université Paris-Saclay. It is a collaboration with portrait photographer Didier Goupy. The installation is designed to be exhibited in various sites of Université Paris-Saclay until it is permanently installed in the Learning Center of Université Paris-Saclay. This project supported Shubhangi Gupta, an intern, for two months over the summer.

8.2. National Initiatives

8.2.1. Investissements d'Avenir

8.2.1.1. Digiscope - Collaborative Interaction with Complex Data and Computation

Type: EQUIPEX (Equipement d'Excellence)

Duration: 2011-2021

Coordinator: Michel Beaudouin-Lafon

Partners: FCS Paris-Saclay (coordinator), Université Paris-Sud, CNRS, CEA, Inria, Institut Mines-Telecom, Ecole Centrale Paris, Université Versailles - Saint-Quentin, ENS Cachan, Maison de la Simulation

Overall budget: 22.5 Meuros, including 6.7 Meuros public funding from ANR

Abstract: The goal of the project is to create ten high-end interactive rooms interconnected by high-speed networks and audio-video facilities to support remote collaboration across interactive visualization environments. The equipment will be open to outside users and targets four main application areas: scientific discovery, product lifetime management, decision support for crisis management, and education and training. Digiscope includes the existing WILD room, and funded the WILDER room. ExSitu contributes its expertise in the design and evaluation of advanced interaction techniques and the development of distributed software architectures for interactive systems. At the end of 2017, all ten rooms and the telepresence network are operational. The project was successfully evaluated by an international jury in June, 2017.

8.3. European Initiatives

8.3.1. European Research Council (ERC)

8.3.1.1. Creating Human-Computer Partnerships

Program: ERC Advanced Grant

Project acronym: CREATIV

Project title: Creating Human-Computer Partnerships

Duration: mois année début - mois année fin

Coordinator: Wendy Mackay

Abstract: CREATIV explores how the concept of co-adaptation can revolutionize the design and use of interactive software. Co-adaptation is the parallel phenomenon in which users both adapt their behavior to the system's constraints, learning its power and idiosyncrasies, and appropriate the system for their own needs, often using it in ways unintended by the system designer. A key insight in designing for co-adaptation is that we can encapsulate interactions and treat them as first class objects, called interaction instruments This lets us focus on the specific characteristics of how human users express their intentions, both learning from and controlling the system. By making instruments co-adaptive, we can radically change how people use interactive systems, providing incrementally learnable paths that offer users greater expressive power and mastery of their technology. The initial goal of the CREATIV project is to fundamentally improve the learning and expressive capabilities of advanced users of creative software, offering significantly enhanced methods for expressing and exploring their ideas. The ultimate goal is to radically transform interactive systems for everyone by creating a powerful and flexible partnership between human users and interactive technology.

8.3.1.2. Unified Principles of Interaction

Program: ERC Advanced Grant

Project acronym: ONE

Project title: Unified Principles of Interaction

Duration: October 2016 - September 2020

Coordinator: Michel Beaudouin-Lafon

Abstract: The goal of ONE is to fundamentally re-think the basic principles and conceptual model of interactive systems to empower users by letting them appropriate their digital environment. The project addresses this challenge through three interleaved strands: empirical studies to better understand interaction in both the physical and digital worlds, theoretical work to create a conceptual model of interaction and interactive systems, and prototype development to test these principles and concepts in the lab and in the field. Drawing inspiration from physics, biology and psychology, the conceptual model combines *substrates* to manage digital information at various levels of abstraction and representation, *instruments* to manipulate substrates, and *environments* to organize substrates and instruments into digital workspaces.

8.4. International Initiatives

8.4.1. Inria Associate Teams Not Involved in an Inria International Labs

8.4.1.1. DECibel

Title: Discover, Express, Create – Interaction Technologies For Creative Collaboration

International Partner (Institution - Laboratory - Researcher):

> University of California Berkeley (United States) - Electrical and Computer Engineering, Center for Magnetic Resonance Research - Bjoern Hartmann

Start year: 2016

See also: https://www.inria.fr/en/associate-team/decibel

The DECibel associated team includes Inria's ExSitu and the CITRIS Connected Communities Initiative (CCI) at UC Berkeley. ExSitu explores extreme interaction, working with creative professionals and scientists who push the limits of technology to develop novel interactive technologies that offer new strategies for creative exploration. ExSitu's research activities include: developing underlying theory (co-adaptive instruments and substrates), conducting empirical studies (participatory design with creative professionals), and implementing interactive systems (creativity support tools). The CITRIS Connected Communities Initiative investigates collaborative discovery and design through new technologies that enhance education, creative work, and public engagement. It develops interactive tools, techniques and materials for the rapid design and prototyping of novel interactive products, expertise sharing among designers, and citizen science investigations. DECibel will combine the strengths of these two groups to to investigate novel tools and technologies that support Discovery, Expressivity, and Creativity.

8.5. International Research Visitors

8.5.1. Visits of International Scientists

- Mobile Life research team (KTH, Sweden) 20 researchers visited ExSitu in January, 2017.
- Susanne Bødker (Aarhus University, Denmark) visited ExSitu in April, 2017.
- Joanna McGrenere (University of British Columbia, Canada) Inria Chair, visited ExSitu in June-July, 2017.

8.5.1.1. Internships

- Alessandro Silacci, Haute Ecole d'Ingenierie et d'Architecture de Fribourg (Suisse), "Cross-Surface Expressive Gesture Interactions in Collaboration Scenarios": Michel Beaudouin-Lafon
- Shubhangi Gupta, "Design and Prototyping of Web Interface to the 'Plateau des Recherches Infinies' Installation": Michel Beaudouin-Lafon
- Alexander Eiselmayer, University of Zurich, "Touchstone II": Wendy Mackay and Michel Beaudouin-Lafon

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. Member of the Organizing Committees

- ACM CHI 2017, *Video-Previews* Chair: Carla Griggio
- ACM CHI 2018, *ACM Conference on Human Factors in Computing*, Web Co-Chair: Ignacio Avellino

9.1.2. Scientific Events Selection

9.1.2.1. Chair of Conference Program Committees

- ACM CHI 2017, *Late-Breaking Work* Co-Chair: Theophanis Tsandilas
- ACM/ISMIR 2018, *Music and Technology* Co-Chair: Wendy Mackay

9.1.2.2. Member of the Conference Program Committees

- ACM CHI 2018, *ACM CHI Conference on Human Factors in Computing Systems*: Michel Beaudouin-Lafon, Wendy Mackay, Theophanis Tsandilas

- ACM CHI 2017, *ACM CHI Conference on Human Factors in Computing Systems*: Michel Beaudouin-Lafon, Wendy Mackay
- ACM UIST 2017, *ACM Symposium on User Interface Software and Technology*: Michel Beaudouin-Lafon, Wendy Mackay
- IUI 2018, *ACM IUI Conference on Intelligent User Interfaces*: Wendy Mackay
- MOCO 2017, *International Conference on Movement and Computing*: Sarah Fdili Alaoui
- EuroVR 2017, *European Virtual Reality Conference*: Cédric Fleury
- ACM C&C 2017, *Creativity and Cognition*: Andrew Webb

9.1.2.3. Reviewer

- ACM CHI 2017-18, *ACM CHI Conference on Human Factors in Computing Systems*: Andrew Webb, Carla Griggio, Nolwenn Maudet, Wanyu Liu, Sarah Fdili Alaoui, Cédric Fleury, Theophanis Tsandilas, Michael Wessely, Ignacio Avellino
- ACM UIST 2017 *ACM Symposium on User Interface Software and Technology*: Andrew Webb, Marianela Ciolfi Felice, Theophanis Tsandilas, Michael Wessely, Wanyu Liu, Nolwenn Maudet, Sarah Fdili Alaoui, Cédric Fleury
- ACM DIS 2017, *ACM Conference on Designing Interactive Systems*: Nolwenn Maudet, Sarah Fdili Alaoui, Wendy Mackay
- ACM ISS 2017, *ACM International Conference on Interactive Surfaces and Spaces*: Ignacio Avellino, Philip Tchernavskij
- IEEE InfoVis 2017: Theophanis Tsandilas
- IEEE VR 2018, *Virtual Reality Conference*: Cédric Fleury
- IEEE 3DUI 2017, *Symposium on 3D User Interfaces*: Cédric Fleury
- GI 2017, *Graphics Interface Conference*: Theophanis Tsandilas
- Interaction 18: Carla Griggio, Germán Leiva
- ACM C&C 2017, *Creativity and Cognition*: Sarah Fdili Alaoui
- Mobile HCI '17 Late Breaking, *19th International Conference on Human-Computer Interaction with Mobile Devices and Services*: Germán Leiva, Philip Tchernavskij
- IHM 2017, *Conférence Francophone d'Interaction Homme-Machine*: Theophanis Tsandilas, Cédric Fleury
- Salon des Refusés 2017, workshop at <PROGRAMMING> 2017, *The International Conference on the Art, Science, and Engineering of Programming*: Philip Tchernavskij

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- Editor for the Human-Computer Interaction area of the ACM Books Series (published with Morgan & Claypool Publishers): Michel Beaudouin-Lafon (2013-)
- CACM, *Communications of the ACM Web Editorial Board*, ACM: Wendy Mackay (2008-)
- ACM, *ACM New Publications Board*, ACM: Wendy Mackay (2015-)
- TOCHI, *Transactions on Computer Human Interaction*, ACM: Michel Beaudouin-Lafon (2009-), Wendy Mackay (2016-)
- JIPS, *Journal d'Interaction Personne-Système*, AFIHM: Michel Beaudouin-Lafon (2009-)

9.1.3.2. Reviewer - Reviewing Activities

- IJHCS, *International Journal of Human-Computer Studies*, Elsevier: Theophanis Tsandilas
- TVCG, *Transactions on Visualization and Computer Graphics*, IEEE: Cédric Fleury
- MTI, *Journal of Multimodal Technologies and Interaction*, MDPI: Sarah Fdili Alaoui

- JHT, *Journal Human Technology, special issue "the Human–Technology Choreographies: Body, movement, and space in expressive interactions"*, EBSCO Publishing: Sarah Fdili Alaoui
- JEP, *Journal of Experimental Psychology*: Michel Beaudouin-Lafon

9.1.4. Invited Talks

- Ecole Normale Supérieur de Lyon, Cultures numériques, éducation aux médias et à l'information, "*Créer des Partenariats Humaine-Machine*", 09 January 2017: Wendy Mackay
- Lyon Culture Numériques Table Ronde, "*Créer des Partenariats Humaine-Machine*", 11 January 2017: Wendy Mackay
- Invited conference, World Information Architecture Day (WIAD), Lyon, "*Interfaces du futur : Unifier pour diversifier*", 18 février 2017: Michel Beaudouin-Lafon
- Invited seminar, CIRMMT (Montreal, Canada), "*Crafting movement in Digital Art and Human Computer Interaction*", 22 March 2017: Sarah Fdili Alaoui
- Colloquim d'Informatique UPMC Sorbonne Universités, Paris, "*Interfaces Homme-Machine : Unifier les principes pour diversifier l'interaction*", 28 February 2017: Michel Beaudouin-Lafon
- Inria 50 years, ERC 10 years, "*ERC CREATIV : Créer des Partenariats Humain-Machine*", 13 March 2017: Wendy Mackay
- IFIP TC13 & TC14 Open symposium (Paris) *Crafting movement in Digital Art and Human Computer Interaction*, 22 March 2017: Sarah Fdili Alaoui
- Wellesley College (Wellesley, MA, USA), "*Towards a Unified Theory of Interaction*", 20 April 2017: Wendy Mackay & Michel Beaudouin-Lafon
- Massachusetts Institute of Technology (Cambridge, MA, USA), "*Towards a Unified Theory of Interaction*", 20 April 2017: Wendy Mackay
- 10Pines (Buenos Aires, Argentina), "*Design and Development of Interactive Systems*", 20 April 2017: Germán Leiva
- 10Pines (Buenos Aires, Argentina), "*Designing cross-application instruments for supporting user appropriation*", 20 April 2017: Carla Griggio
- Participatory Design Workshop, Olin College (Needham, MA, USA), "*Apps for Activists?*", 21 April 2017: Wendy Mackay
- CITEP UBA (Buenos Aires, Argentina) "*Introduction to HCI*", 25 April October 2017: Germán Leiva
- CITEP UBA (Buenos Aires, Argentina) "*Designing cross-application instruments for supporting user appropriation*", 25 April October 2017: Carla Griggio
- Atelier Unibail-Rodamco, "*Créer des Partenariats Humain-Machine*", 02 May 2017: Wendy Mackay
- CHI Stories, Invited Address, ACM CHI'17, "*Creating the First Interactive Video*", 09 May 2017: Wendy Mackay
- Invited Address, University of British Columbia, "*Towards Unified Principles of Interaction*", 09 May 2017: Wendy Mackay
- Dagstuhl Seminar, "*Computational Interactivity*", 5 - 8 June 2017: Wanyu Liu (http://www.dagstuhl.de/en/program/calendar/semhp/?semnr=17232)
- Colloque 10 ans de l'ERC, Université Paris-Saclay, "ONE: Unified Principles of Interaction", 9 June 2017: Michel Beaudouin-Lafon
- ERC CREATIVE Workshop, "*Towards a Unified Theory of Interaction*", 13 June 2017: Wendy Mackay
- University of British Columbia (Canada), "*Towards Unified Principles of Interaction*", 15 June 2017: Michel Beaudouin-Lafon

- Rencontres Jeunes Cherchers, Interaction Homme-Machine, "*Ethics and Experimental Protocols*", 04 July 2017: Wendy Mackay
- Participatory Information Technology group, Aarhus University, "*Untangling Interactive Systems*", 14 September 2017: Philip Tchernavskij
- Honoris Causa Invited Address, Aarhus University (Denmark), "*Human-Computer Partnerships*", 14 September 2017: Wendy Mackay
- Aarhus University (Denmark), "*Towards Unified Principles of Interaction*", 14 September 2017: Michel Beaudouin-Lafon
- Opening Keynote, 12th Biannual Conference of the Italian SIGCHI Chapter (CHItaly 17), "*Towards Unified Principles of Interaction*", 19 September 2017: Michel Beaudouin-Lafon
- Journées sur la Dynamique Interactionnelle du Geste DIG, CNRS, *Applying Laban Movement Analysis to craft movement in Digital Art and Human Computer Interaction*, 22 septembre 2017: Sarah Fdili Alaoui
- Journée LTCI , "*Bayesian Information Gain for Multiscale Navigation*", 10 October 2017: Wanyu Liu
- Les object communicants dans l'écologie des enfants de 0 à 12 ans, MENESR-EHESS, "*Entre les enfants et les ordinateurs: Human Computer Interaction*", 11 October 2017: Wendy Mackay
- ETH Zurich and University of Zurich, "*Understanding and Designing Interaction from an Information-Theoretic Perspective*", 24 November 2017: Wanyu Liu
- Inria 50 Years Celebration, Robotics panel, 8 November 2017: Wendy Mackay
- Remise des Prix Doctorants, "*Bayesian Information Gain for Multiscale Navigation*", 30 November 2017: Wanyu Liu
- Colloque "Corps et Mobiles", Sorbonne Nouvelle, *Radical Choreographic Object*, 4-5 December 2017: Sarah Fdili Alaoui
- Meeting on Global Research Network, Casa Paganini Infomuse, Genova, 2017: Sarah Fdili Alaoui

9.1.5. Scientific Expertise

- European Research Council, Panel member for Starting Grants: Michel Beaudouin-Lafon
- CNRS Mission pour l'Interdisciplinarité, Panel member for the call "Sciences Sociales et Cognitives des Comportements Collectifs": Michel Beaudouin-Lafon
- CNRS INS2I, member of "Cellule ERC": Michel Beaudouin-Lafon
- ACM "Policy Award" committee member: Michel Beaudouin-Lafon
- ACM SIGCHI "Lifetime Service Award" committee member: Wendy Mackay

9.1.6. Research Administration

- CNRS INS2I, "Conseil Scientifique de l'Institut": Michel Beaudouin-Lafon (member)
- Telecom ParisTech "Comité de la recherche": Michel Beaudouin-Lafon (member)
- IRCAM Scientific Committee: Michel Beaudouin-Lafon (member)
- "Institut de la Société Numérique", IDEX Laboratory of Université Paris-Saclay: Michel Beaudouin-Lafon (member of steering committee)
- Pôle Systematic, Working group on Information Systems: Michel Beaudouin-Lafon (member of steering committee)
- Département STIC, Université Paris-Saclay: Wendy Mackay (member)
- "Conseil de Laboratoire", LRI: Wendy Mackay, Cédric Fleury (members)
- "Conseil Scientifique", LRI: Michel Beaudouin-Lafon (member)

- COERLE “Comité Operationnel d’Evaluation des Risques Légaux et Ethiques”, Inria: Wendy Mackay
- CERNI “Comité d’Ethique pour les Recherches Non Interventionnelles”, Université Paris-Sud: Wendy Mackay
- CCSU, “Commission Consultative de Spécialistes de l’Université”, Université Paris-Sud: Michel Beaudouin-Lafon, Wendy Mackay (members)
- “Commission de Qualification pour Promotion et Changement, Professeur”, Télécom Paris-Tech: Wendy Mackay (member)
- “Comité de sélection, Professeur”, Université Paris-Sud: Wendy Mackay (member)
- “Comité de sélection, Maître de Conférences”, Université Paris-Sud: Michel Beaudouin-Lafon (member)
- “Commission Locaux”, LRI: Theophanis Tsandilas (member)
- “Commission Scientifique”, Inria: Theophanis Tsanidlas (member), since March 2017
- “User board” of the European project Wholodance: Sarah Fdili Alaoui, since June 2017

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence : Sarah Fdili Alaoui, *Programmation des interfaces interactives avancées*, 22.5h, L3, Univ. Paris-Sud

International Masters: Theophanis Tsandilas, *Probabilities and Statistics*, 32h, M1, Univ. Paris-Saclay

HCID Masters: Sarah Fdili Alaoui, *Business Development Labs*, 30h, M1, Univ. Paris-Sud

HCID Masters: Sarah Fdili Alaoui, *Innovation & Entrepreneurship thesis*, 3h, M2, Univ. Paris-Sud

HCID Masters: Sarah Fdili Alaoui, *Design Project*, 36h, M1 et M2, Univ. Paris-Sud

HCID Masters: Michel Beaudouin-Lafon, Wendy Mackay, *Fundamentals of Situated Interaction*, 21 hrs, M1/M2, Univ. Paris-Sud

Interaction & HCID Masters: Sarah Fdili Alaoui, *Stage en entreprise*, 2h, M2, Univ. Paris-Sud

Interaction & HCID Masters: Sarah Fdili Alaoui, *Creative Design*, 27h, M1 et M2, Univ. Paris-Sud

Interaction & HCID Masters: Sarah Fdili Alaoui, *Digital Fabrication*, 13,5h, M1 et M2, Univ. Paris-Sud

Interaction & HCID Masters: Michel Beaudouin-Lafon, *Fundamentals of Human-Computer Interaction*, 21 hrs, M1/M2, Univ. Paris-Sud

Interaction & HCID Masters: Michel Beaudouin-Lafon & Cédric Fleury, *Groupware and Collaborative Interaction*, 31.5 hrs, M1/M2, Univ. Paris-Sud

Interaction & HCID Masters: Wendy Mackay, *Career Seminar* 6 hrs, M2, Univ. Paris-Sud

Interaction & HCID Masters: Wendy Mackay, *Design of Interactive Systems*, 21 hrs, M1/M2, Univ. Paris-Sud

Interaction & HCID Masters: Wendy Mackay, *Advanced Design of Interactive Systems*, 21 hrs, M1/M2, Univ. Paris-Sud

Polytech: Sarah Fdili Alaoui, *Graphisme et Visualisation*, 18h, “Apprentis” 5th year, Univ. Paris-Sud

Polytech: Cédric Fleury, *Introduction à l’Informatique*, 71 hrs, 1st year, Univ. Paris-Sud

Polytech: Cédric Fleury, *Projet Java-Graphique-IHM*, 24 hrs, 3st year, Univ. Paris-Sud

Polytech: Cédric Fleury, *Interaction Homme-Machine*, 18 hrs, 3st year, Univ. Paris-Sud

Polytech: Cédric Fleury, *Réalité Virtuelle et Interaction*, 48 hrs, 5th year, Univ. Paris-Sud

9.2.2. Supervision

PhD students

PhD: Ignacio Avellino, *Supporting Collaborative Practices Across Wall-Sized Displays with Video-Mediated Communication*, Université Paris-Saclay, 12 December 2017. Advisors: Michel Beaudouin-Lafon & Cédric Fleury

PhD: Jessalyn Alvina, *Increasing The Expressive Power of Gesture-based Interaction on Mobile Devices*, Université Paris-Saclay, 13 December 2017. Advisor: Wendy Mackay

PhD: Nolwenn Maudet, *Designing Design Tools*, Université Paris-Saclay, 11 December 2017. Advisors: Wendy Mackay & Michel Beaudouin-Lafon

PhD in progress: Marianela Ciolfi Felice, *Substrates and Co-adaptive Instruments for Creativity*, September 2015. Advisors: Wendy Mackay & Sarah Fdili Alaoui

PhD in progress: Carla Griggio, *Interactive Human-Machine Learning*, September 2015. Advisor: Wendy Mackay

PhD in progress: Germán Leiva, *Interaction-driven Software Development*, September 2015. Advisor: Michel Beaudouin-Lafon

PhD in progress: Wanyu Liu, *Modeling the speed-accuracy trade-off of pointing tasks using the tools of information theory*, October 2015. Advisors: Olivier Rioul (Institut Mines Telecom) & Michel Beaudouin-Lafon

PhD in progress: Yujiro Okuya, *Sensorimotor interface for Collaborative Virtual Environments based on heterogeneous interactive devices: application to industrial design*, October 2015. Advisors: Patrick Bourdot (LIMSI-CNRS) & Cédric Fleury

PhD in progress: Michael Wessely, *Sketching and Physical Prototyping for Creative Fabrication Design*, November 2015. Advisors: Theophanis Tsandilas & Wendy Mackay

PhD in progress: Stacy (Shu-Yuan) Hsueh, *Embodied design for Human-Computer Co-creation*, November 2015. Advisors: Wendy Mackay & Sarah Fdili Alaoui

PhD in progress: Philip Tchernavskij, *Towards Unified Principles of Interaction*, October 2016. Advisor: Michel Beaudouin-Lafon

PhD in progress: Jean-Philippe Rivière, *Embodied Design for Human-Computer Partnership in Learning Contexts*, October 2017. Advisors: Wendy Mackay & Sarah Fdili Alaoui

Masters students

Jiali Liu, Telecom ParisTech, “Cross-Surface Interaction”: Michel Beaudouin-Lafon & Philip Tchernavskij

Linghua Lai, Univ. Paris-Saclay, “VideoBoard: Video for Interaction Design”: Wendy Mackay & Germán Leiva

Chengcheng Qu, Univ. Paris-Saclay, “Animated Emojis: Creating Dynamic Expressions with Expressive Keyboards”: Wendy Mackay & Jessalyn Alvina

Jean-Philippe Rivière, Univ. Toulouse, “MoveOn: Understanding and Supporting Dance Skill Acquisition through Multimodal Interactive Technology”: Wendy Mackay & Sarah Fdili Alaoui

9.2.3. Juries

PhD theses:

- Andrew Bluff, University of Technology Sydney (advisor: Andrew Johnston): Sarah Fdili Alaoui.
- Bart Potsma, LIMSI-CNRS, April 2017 (advisor: Brian Katz): Michel Beaudouin-Lafon, president.
- Samuel Delalez, Université Paris-Saclay, November 2017 (advisor: Christophe d’Alessandro): Wendy Mackay, president
- Maxime Guillon, LIG, November 2017 (advisor: Laurence Nigay): Michel Beaudouin-Lafon, reviewer.
- Hèlene Unaninski, Université de Toulouse, December 2017 (advisor: Stéphane Conversy): Wendy Mackay, reviewer

Habilitations:

- Caroline Appert, LRI, juin 2017: Michel Beaudouin-Lafon, examiner.
- Gilles Bailly, ISIR, octobre 2017: Michel Beaudouin-Lafon, examiner.
- Jean-Julien Aucouturier, IRCAM, novembre 2017: Michel Beaudouin-Lafon, president.

9.3. Popularization

- Interviews on France Culture, LCI, TV5 Monde on the standardization of the AZERTY and BEPO keyboards, June-July 2017, Michel Beaudouin-Lafon
- ACM UIST 17 Demonstration, "*CommandBoard: Creating a General-Purpose Command Gesture Input Space for Soft Keyboard*", Jessalyn Alvina, Carla Griggio, Xiaojun Bi, Wendy Mackay
- Interview à France Culture, 22 Février 2017 @ France Culture *Machines au croisement du design et du numérique*: Nolwenn Maudet
- Demonstration at CEA Digital Days: Data Intelligence, CEA List, "*StickyLines*", March 2017, Marianela Ciolfi Felice & Wendy Mackay
- Demonstration at CEA Digital Days: Data Intelligence, CEA List, "*Expressive Keyboard*", March 2017, Jessalyn Alvina & Wendy Mackay
- ACM SCF 2017 Poster Demonstration *Stretchis: Fabricating Highly Stretchable User Interfaces*: Michael Wessely
- Colloque l'Université Paris-Saclay fête les 10 ans de l'ERC, "*ONE - Unified Principles of Interaction*", 9 June 2017: Michel Beaudouin-Lafon (presentation) and Philip Tchernavskij (poster)
- Presentation of the performance *Radical Choreographic Object*, April 2017 in Theatre de Villeneuve tolosane, September 2017 in Centre de développement Choregraphique, October 2017 in Centre Bellegarde, November 2017 in Le 104 and December 2017 in Centre Bellegarde: Sarah Fdili Alaoui
- Presentation of the Performance *SKIN*, April 2017 in Scene 44 Pole Media, La friche de la belle, May 2017 in festival Curiositas, November 2017 in Attention Travaux, Culture en Essonnes: Sarah Fdili Alaoui
- Scaling up for the Internet of Things, Workshop at Inria Silicon Valley, Berkeley, CA, USA: Wendy Mackay & Theophanis Tsandilas
- Intersections in Music, Movement, and Technology, 17–19 August 2017 *Workshop held at UC Berkeley*: John MacCallum (co-facilitator).
- Relational Listening, 8 July 2017 *Workshop presented at the Conference of Dance and Somatic Practices*: John MacCallum (co-facilitator).
- Relational Listening, 1–30 April 2017 *Workshop*: John MacCallum (co-facilitator).
- Relational Listening, 26 February 2017 *Workshop held at CLOUD/Danslab*: John MacCallum (co-facilitator).

10. Bibliography

Major publications by the team in recent years

[1] J. Alvina, J. Malloch, W. Mackay.*Expressive Keyboards: Enriching Gesture-Typing on Mobile Devices*, in "Proceedings of the 29th ACM Symposium on User Interface Software and Technology (UIST 2016)", Tokyo, Japan, ACM (editor), ACM, October 2016, p. 583 - 593 [*DOI :* 10.1145/2984511.2984560], https://hal.inria.fr/hal-01437054.

[2] A. BOUSSEAU, T. TSANDILAS, L. OEHLBERG, W. E. MACKAY.*How Novices Sketch and Prototype Hand-Fabricated Objects*, in "Conference on Human Factors in Computing Systems (CHI)", San Jose, United States, ACM, May 2016 [*DOI :* 10.1145/2858036.2858159], https://hal.inria.fr/hal-01272187.

[3] M. CIOLFI FELICE, N. MAUDET, W. E. MACKAY, M. BEAUDOUIN-LAFON.*Beyond Snapping: Persistent, Tweakable Alignment and Distribution with StickyLines*, in "UIST '16 Proceedings of the 29th Annual Symposium on User Interface Software and Technology", Tokyo, Japan, Proceedings of the 29th Annual Symposium on User Interface Software and Technology, October 2016 [*DOI :* 10.1145/2984511.2984577], https://hal.archives-ouvertes.fr/hal-01410171.

[4] G. JALAL, N. MAUDET, W. E. MACKAY.*Color Portraits: From Color Picking to Interacting with Color*, in "Proceedings of the ACM International Conference on Human Factors in Computing Systems, 2015", Seoul, South Korea, May 2015 [*DOI :* 10.1145/2702123.2702173], https://hal.archives-ouvertes.fr/hal-01226494.

[5] C. KLOKMOSE, J. EAGAN, S. BAADER, W. MACKAY, M. BEAUDOUIN-LAFON.*Webstrates: Shareable Dynamic Media*, in "ACM Symposium on User Interface Software and Technology (UIST)", Charlotte, United States, ACM (editor), UIST '15 Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology, November 2015, p. 280-290 [*DOI :* 10.1145/2807442.2807446], https://hal.archives-ouvertes.fr/hal-01242672.

[6] C. LIU, O. CHAPUIS, M. BEAUDOUIN-LAFON, E. LECOLINET.*Shared Interaction on a Wall-Sized Display in a Data Manipulation Task*, in "Proceedings of the 34th international conference on Human factors in computing systems", San Jose, United States, ACM (editor), CHI '16, SigCHI and ACM, May 2016, p. 1-12 [*DOI :* 10.1145/2858036.2858039], https://hal.inria.fr/hal-01275535.

[7] J.-B. LOUVET, C. FLEURY.*Combining Bimanual Interaction and Teleportation for 3D Manipulation on Multi-Touch Wall-sized Displays*, in "Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology (VRST)", Munich, Germany, November 2016, p. 283 - 292 [*DOI :* 10.1145/2993369.2993390], https://hal.archives-ouvertes.fr/hal-01428914.

[8] M. NANCEL, E. PIETRIGA, O. CHAPUIS, M. BEAUDOUIN-LAFON.*Mid-air Pointing on Ultra-Walls*, in "ACM Transactions on Computer-Human Interaction", August 2015, vol. 22, n^{o} 5, p. 21:1–21:62 [*DOI :* 10.1145/2766448], https://hal.inria.fr/hal-01184544.

[9] T. TSANDILAS, A. BEZERIANOS, T. JACOB.*SketchSliders: Sketching Widgets for Visual Exploration on Wall Displays*, in "Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems", Seoul, South Korea, ACM, April 2015, p. 3255-3264 [*DOI :* 10.1145/2702123.2702129], https://hal.archives-ouvertes.fr/hal-01144312.

[10] M. WESSELY, T. TSANDILAS, W. E. MACKAY.*Stretchis: Fabricating Highly Stretchable User Interfaces*, in "ACM Symposium on User Interface Software and Technology (UIST)", Tokyo, Japan, October 2016, p. 697-704 [*DOI :* 10.1145/2984511.2984521], https://hal.archives-ouvertes.fr/hal-01412289.

Publications of the year

Articles in International Peer-Reviewed Journal

[11] S. FDILI ALAOUI.*Vers une dramaturgie et un corps au numérique*, in "Revue Culture et Recherche", February 2017, vol. 135, 1, https://hal.inria.fr/hal-01680445.

International Conferences with Proceedings

[12] S. F. ALAOUI.*Workshop on Experience Explicitation Interviews for Movement Researchers and Practitioners*, in "MOCO 2017 - 4th International conference on movement and computing", Londres, United Kingdom, June 2017, https://hal.inria.fr/hal-01680450.

[13] S. F. ALAOUI, J. FRANÇOISE, T. SCHIPHORST, K. STUDD, F. BEVILACQUA.*Seeing, Sensing and Recognizing Laban Movement Qualities*, in "ACM Conference on Human Factors in Computing Systems (CHI)", Denver, United States, May 2017 [*DOI :* 10.1145/3025453.3025530], https://hal.inria.fr/hal-01663132.

[14] J. ALVINA, C. F. GRIGGIO, X. BI, W. E. MACKAY.*CommandBoard: Creating a General-Purpose Command Gesture Input Space for Soft Keyboards*, in "UIST '17 Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology", Quebec City, Canada, October 2017, p. 17-28 [*DOI :* 10.1145/3126594.3126639], https://hal.archives-ouvertes.fr/hal-01679137.

[15] I. AVELLINO, C. FLEURY, W. MACKAY, M. BEAUDOUIN-LAFON.*CamRay: Camera Arrays Support Remote Collaboration on Wall-Sized Displays*, in "Proceedings of the CHI Conference on Human Factors in Computing Systems", Denver, United States, ACM (editor), CHI '17, May 2017, p. 6718 - 6729 [*DOI :* 10.1145/3025453.3025604], https://hal.archives-ouvertes.fr/hal-01544645.

[16] Y. CANDAU, J. FRANÇOISE, S. F. ALAOUI, T. SCHIPHORST.*Cultivating kinaesthetic awareness through interaction Perspectives from somatic practices and embodied cognition*, in "MOCO 2017 - 4th International Conference on Movement Computing", Londres, United Kingdom, June 2017, p. 1-8 [*DOI :* 10.1145/3077981.3078042], https://hal.inria.fr/hal-01680387.

[17] J. FRANÇOISE, Y. CANDAU, S. F. ALAOUI, T. SCHIPHORST.*Designing for Kinesthetic Awareness: Revealing User Experiences through Second-Person Inquiry*, in "CHI 2017 - ACM Conference on Human Factors in Computing Systems", Denver, United States, ACM, May 2017, p. 5171-5183 [*DOI :* 10.1145/3025453.3025714], https://hal.inria.fr/hal-01662526.

[18] J. GORI, O. RIOUL, Y. GUIARD, M. BEAUDOUIN-LAFON.*One Fitts' Law, Two Metrics*, in "INTERACT 2017 - 16th IFIP TC.13 International Conference on Human-Computer Interaction", Mumbai, India, September 2017 [*DOI :* 10.1007/978-3-319-67687-6_36], https://hal.archives-ouvertes.fr/hal-01675635.

[19] C. LIU, O. CHAPUIS, M. BEAUDOUIN-LAFON, E. LECOLINET.*CoReach: Cooperative Gestures for Data Manipulation on Wall-sized Displays*, in "Proceedings of the 35th international conference on Human factors in computing systems", Denver, United States, ACM (editor), CHI '17, May 2017 [*DOI :* 10.1145/3025453.3025594], https://hal.archives-ouvertes.fr/hal-01437091.

[20] *Best Paper*
W. LIU, R. LUCAS D'OLIVEIRA, M. BEAUDOUIN-LAFON, O. RIOUL.*BIGnav: Bayesian Information Gain for Guiding Multiscale Navigation*, in "ACM CHI 2017 - International conference of Human-Computer Interaction", Denver, United States, May 2017, p. 5869-5880 [*DOI :* 10.1145/3025453.3025524], https://hal.inria.fr/hal-01677122.

[21] W. LIU, O. RIOUL, M. BEAUDOUIN-LAFON, Y. GUIARD.*Information-Theoretic Analysis of Human Performance for Command Selection*, in "INTERACT 2017 - 16th IFIP TC 13 International Conference on

Human-Computer Interaction", Mumbai, India, LNCS, Springer, September 2017, vol. 10515, p. 515-524 [*DOI :* 10.1007/978-3-319-67687-6_35], https://hal.inria.fr/hal-01643924.

[22] J. MALLOCH, C. F. GRIGGIO, J. MCGRENERE, W. MACKAY.*Fieldward and Pathward: Dynamic Guides for Defining Your Own Gestures*, in "CHI '17 Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems", Denver, United States, May 2017, vol. 55, n^{o} 9, p. 4266 - 4277 [*DOI :* 10.1145/3025453.3025764], https://hal.inria.fr/hal-01614267.

[23] N. MAUDET, G. JALAL, P. TCHERNAVSKIJ, M. BEAUDOUIN-LAFON, W. MACKAY.*Beyond Grids, Interactive Graphical Substrates to Structure Digital Layout*, in "Proceedings of the ACM International Conference on Human Factors in Computing Systems, 2017", Denver, United States, ACM (editor), ACM, May 2017, p. 5053 - 5064 [*DOI :* 10.1145/3025453.3025718], https://hal.archives-ouvertes.fr/hal-01609030.

[24] N. MAUDET, G. LEIVA, M. BEAUDOUIN-LAFON, W. E. MACKAY.*Design Breakdowns: Designer-Developer Gaps in Representing and Interpreting Interactive Systems*, in "CSCW 2017 - ACM Conference on Computer Supported Cooperative Work and Social Computing", Portland, Oregon, United States, February 2017, p. 630 - 641 [*DOI :* 10.1145/2998181.2998190], https://hal.inria.fr/hal-01614250.

[25] T. J. NACCARATO, J. MACCALLUM.*Critical Appropriations of Biosensors in Artistic Practice*, in "MOCO 2017 - 4th International conference on movement and computing", London, United Kingdom, June 2017 [*DOI :* 10.1145/3077981.3078053], https://hal.inria.fr/hal-01614245.

[26] M. NOUWENS, C. F. GRIGGIO, W. E. MACKAY.*"WhatsApp is for family; Messenger is for friends": Communication Places in App Ecosystems*, in "CHI '17 - Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems", Denver, United States, May 2017, p. 727 - 735 [*DOI :* 10.1145/3025453.3025484], https://hal.inria.fr/hal-01614125.

[27] M. PIOVARČI, M. WESSELY, M. JAGIELSKI, M. ALEXA, W. MATUSIK, P. DIDYK.*Directional Screens*, in "SCF '17 - Proceedings of the 1st Annual ACM Symposium on Computational Fabrication", Cambridge, United States, June 2017, p. 1 - 10 [*DOI :* 10.1145/3083157.3083162], https://hal.archives-ouvertes.fr/hal-01626159.

[28] P. TCHERNAVSKIJ, C. N. KLOKMOSE, M. BEAUDOUIN-LAFON.*What Can Software Learn From Hypermedia?*, in "Programming '17 Companion to the first International Conference on the Art, Science and Engineering of Programming", Brussels, Belgium, ACM, April 2017 [*DOI :* 10.1145/3079368.3079408], https://hal.inria.fr/hal-01614254.

[29] P. TCHERNAVSKIJ.*Decomposing Interactive Systems*, in "CHI 2017 - Tools workshop at CHI'2017", Denver, United States, May 2017, 4, https://hal.inria.fr/hal-01614246.

[30] F. VITALE, J. MCGRENERE, A. TABARD, M. BEAUDOUIN-LAFON, W. MACKAY.*High Costs and Small Benefits: A Field Study of How Users Experience Operating System Upgrades*, in "CHI 2017", Denver, United States, Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, ACM, May 2017, p. 4242-4253 [*DOI :* 10.1145/3025453.3025509], https://hal.archives-ouvertes.fr/hal-01493415.

[31] J. WANG, J. MALLOCH, S. HUOT, F. CHEVALIER, M. WANDERLEY.*Versioning and Annotation Support for Collaborative Mapping Design*, in "Sound and Music Computing Conference", Espoo, Finland, Aalto University, July 2017, p. 275-278, https://hal.inria.fr/hal-01576143.

Conferences without Proceedings

[32] K. B. ANTONSEN, M. BEAUDOUIN-LAFON, J. EAGAN, C. N. KLOKMOSE, W. MACKAY, R. RÄDLE.*Webstrates for the Future Web?*, in "ProWeb 2017 - Programming Technology for the Future Web", Brussels, Belgium, April 2017, 1, https://hal.inria.fr/hal-01614236.

[33] M. BEAUDOUIN-LAFON.*Towards Unified Principles of Interaction*, in "CHItaly 2017 - Proceedings of the 12th Biannual Conference of the Italian SIGCHI Chapter", Cagliari, Italy, September 2017, p. 1 - 2 [*DOI :* 10.1145/3125571.3125602], https://hal.archives-ouvertes.fr/hal-01614273.

Scientific Popularization

[34] D. AUBER, M. BEAUDOUIN-LAFON.*Visualisation de données massives*, in "Les Big Data à découvert", M. BOUZEGHOUB, R. MOSSERI (editors), CNRS ÉDITIIONS, March 2017, https://hal.archives-ouvertes.fr/hal-01515255.

[35] J. LI, M. WESSELY, S. FOLLMER, S. MUELLER.*Summer School for Computational Fabrication and Smart Matter*, in "IEEE Pervasive Computing", October 2017, vol. 16, n^o 4, p. 50-53 [*DOI :* 10.1109/MPRV.2017.3971135], https://hal.archives-ouvertes.fr/hal-01675607.

Other Publications

[36] W. LIU, R. LUCAS D'OLIVEIRA, M. BEAUDOUIN-LAFON, O. RIOUL.*A Bayesian Experimental Design Approach Maximizing Information Gain for Human-Computer Interaction*, February 2017, 1, ITA 2017 - IEEE Information Theory and Applications Workshop, Poster, https://hal.inria.fr/hal-01677034.

[37] N. MARQUARDT, S. HOUBEN, M. BEAUDOUIN-LAFON, A. D. WILSON.*HCITools: Strategies and Best Practices for Designing, Evaluating and Sharing Technical HCI Toolkits*, May 2017, p. 624-627, Workshop of the CHI 2017 Conference on Human Factors in Computing Systems [*DOI :* 10.1145/3027063.3027073], https://hal.archives-ouvertes.fr/hal-01614264.

References in notes

[38] J. ALVINA.*Increasing the Expressive Power of Gesture-based Interaction on Mobile Devices*, Université Paris-Saclay, Orsay, France, December 2017.

[39] I. AVELLINO.*Supporting Collaborative Practices across Wall-Sized Displays with Video-Mediated Communication*, Université Paris-Saclay, Orsay, France, December 2017.

[40] N. MAUDET.*Designing Design Tools*, Université Paris-Saclay, Orsay, France, December 2017.

Project-Team GALEN

Organ Modeling through Extraction, Representation and Understanding of Medical Image Content

IN PARTNERSHIP WITH:
Ecole Centrale Paris

RESEARCH CENTER
Saclay - Île-de-France

THEME
Computational Neuroscience and Medicine

Table of contents

Project-Team GALEN

Creation of the Team: 2008 February 15, updated into Project-Team: 2014 January 01, end of the Project-Team: 2017 December 31

Keywords:

Computer Science and Digital Science:

A3.4. - Machine learning and statistics
A3.4.1. - Supervised learning
A3.4.2. - Unsupervised learning
A3.4.3. - Reinforcement learning
A3.4.4. - Optimization and learning
A3.4.5. - Bayesian methods
A3.4.6. - Neural networks
A3.4.7. - Kernel methods
A3.4.8. - Deep learning
A5.3.1. - Compression
A5.3.2. - Sparse modeling and image representation
A5.3.3. - Pattern recognition
A5.3.4. - Registration
A5.4.1. - Object recognition
A5.4.2. - Activity recognition
A5.4.4. - 3D and spatio-temporal reconstruction
A5.9. - Signal processing
A5.9.1. - Sampling, acquisition
A5.9.2. - Estimation, modeling
A5.9.3. - Reconstruction, enhancement
A5.9.4. - Signal processing over graphs
A5.9.5. - Sparsity-aware processing
A5.9.6. - Optimization tools
A6.2.3. - Probabilistic methods
A6.2.4. - Statistical methods
A6.2.6. - Optimization
A6.2.7. - High performance computing
A6.3.1. - Inverse problems
A8.1. - Discrete mathematics, combinatorics
A8.2. - Optimization
A8.6. - Information theory
A8.7. - Graph theory
A9.2. - Machine learning
A9.3. - Signal analysis
A9.5. - Robotics

Other Research Topics and Application Domains:

B2.2.3. - Cancer
B2.6.1. - Brain imaging
B2.6.2. - Cardiac imaging
B2.6.3. - Biological Imaging
B9.4.2. - Mathematics
B9.4.5. - Data science

1. Personnel

Faculty Members

Nikolaos Paragyios [Team leader, Centrale-Supélec, Professor, HDR]
Emilie Chouzenoux [Inria and Univ Paris-Est Marne La Vallée, Associate Professor]
Fragkiskos Malliaros [Centrale-Supélec, Assistant Professor, from Oct 2017]
Jean-Christophe Pesquet [Centrale-Supélec, Professor]

External Collaborators

Iasonas Kokkinos [Facebook France SARL and University College London, Department of Computer Science]
Evangelia Zacharaki [Centrale-Supélec, until Jun 2017. University of Patras, from Jul 2017]

Technical Staff

Norbert Bus [Centrale-Supélec, from Mar 2017]
Natalia Leclercq [Centrale-Supélec, until Aug 2017]
Jana Stern [Centrale-Supélec, from Sep 2017]
Rafael Marini Silva [Centrale-Supélec]

PhD Students

Eugène Belilovsky [Centrale-Supélec]
Maxim Berman [Centrale-Supélec]
Diane Bouchacourt [Centrale-Supélec, until Sep 2017]
Siddhartha Chandra [Inria]
Guillaume Chassagnon [Centrale-Supélec and Hôpitaux Universitaires Paris Centre, Cochin.]
Marie-Caroline Corbineau [Centrale-Supélec]
Riza Guler [Inria]
Hariprasad Kannan [Centrale-Supélec]
Abdulkadir Celikkanat [Centrale-Supélec]
Stefan Kinauer [Inria]
Evgenios Kornaropoulos [Centrale-Supélec, until Feb 2017]
Huu Dien Khue Le [Centrale-Supélec]
Azar Louzi [Centrale-Supélec, until May 2017]
Arthur Marmin [Centrale-Supélec, from Sep 2017]
Mihir Sahasrabudhe [Centrale-Supélec]
Maissa Sghaier [Centrale-Supélec, from Oct 2017]
Jiaqian Yu [Centrale-Supélec, until Feb 2017]

Post-Doctoral Fellow

Maria Vakalopoulou [Centrale-Supélec, from Jun 2017]

Visiting Scientist

Dimitris Samaras [Centrale-Supélec]

Administrative Assistant

Alexandra Merlin [Inria]

2. Overall Objectives

2.1. GALEN@Centrale-Paris

Computational vision is one of the most challenging research domains in engineering sciences. The aim is to reproduce human visual perception through intelligent processing of visual data. The application domains span from computer aided diagnosis to industrial automation & robotics. The most common mathematical formulation to address such a challenge is through mathematical modeling. In such a context, first the solution of the desired vision task is expressed in the form of a parameterized mathematical model. Given such a model, the next task consists of associating the model parameters with the available observations, which is often called the model-to-data association. The aim of this task is to determine the impact of a parameter choice to the observations and eventually maximize/minimize the adequacy of these parameters with the visual observations. In simple words, the better the solution is, the better it will be able to express and fit the data. This is often achieved through the definition of an objective function on the parametric space of the model. Last, but not least given the definition of the objective function, visual perception is addressed through its optimization with respect to the model parameters. To summarize, computation visual perception involves three aspects, a task-specific definition of a parametric model, a data-specific association of this model with the available observations and last the optimization of the model parameters given the objective and the observations.

Such a chain processing inherits important shortcomings. The curse of dimensionality is often used to express the importance of the model complexity. In simple words, the higher the complexity of the model is, the better its expressive power will be with counter effect the increase of the difficulty of the inference process. Non-linearity is another issue to be addressed which simply states that the association between the model and the data is a (highly) non-linear function and therefore direct inference is almost infeasible. The impact of this aspect is enforced from the curse of non-convexity that characterizes the objective function. Often it lives in high-dimensional spaces and is ill posed making exact inference problematic (in many cases not possible) and computationally expensive. Last, but not least modularity and scalability is another important concern to be addressed in the context of computational vision. The use of task-specific modeling and algorithmic solutions make their portability infeasible and therefore transfer of knowledge from one task to another is not straightforward while the methods do not always scale well with respect either to the dimensionality of the representation or the data.

GALEN aims at proposing innovative techniques towards automatic structuring, interpretation and longitudinal modeling of visual data. In order to address these fundamental problems of computational perception, GALEN investigates the use of discrete models of varying complexity. These methods exhibit an important number of strengths such as their ability to be modular with respect to the input measurements (clinical data), the nature of the model (certain constraints are imposed from computational perspective in terms of the level of interactions), and the model-to-data association while being computational efficient.

3. Research Program

3.1. Shape, Grouping and Recognition

A general framework for the fundamental problems of image segmentation, object recognition and scene analysis is the interpretation of an image in terms of a set of symbols and relations among them. Abstractly stated, image interpretation amounts to mapping an observed image, X to a set of symbols Y. Of particular interest are the symbols Y^* that *optimally explain the underlying image*, as measured by a scoring function s that aims at distinguishing correct (consistent with human labellings) from incorrect interpretations:

$$Y^* = \mathrm{argmax}_Y s(X, Y) \tag{1}$$

Applying this framework requires (a) identifying which symbols and relations to use (b) learning a scoring function s from training data and (c) optimizing over Y in Eq.1.

One of the main themes of our work is the development of methods that jointly address (a,b,c) in a shape-grouping framework in order to reliably extract, describe, model and detect shape information from natural and medical images. A principal motivation for using a shape-based framework is the understanding that shape- and more generally, grouping- based representations can go all the way from image features to objects. Regarding aspect (a), image representation, we cater for the extraction of image features that respect the shape properties of image structures. Such features are typically constructed to be purely geometric (e.g. boundaries, symmetry axes, image segments), or appearance-based, such as image descriptors. The use of machine learning has been shown to facilitate the robust and efficient extraction of such features, while the grouping of local evidence is known to be necessary to disambiguate the potentially noisy local measurements. In our research we have worked on improving feature extraction, proposing novel blends of invariant geometric- and appearance- based features, as well as grouping algorithms that allow for the efficient construction of optimal assemblies of local features.

Regarding aspect (b) we have worked on learning scoring functions for detection with deformable models that can exploit the developed low-level representations, while also being amenable to efficient optimization. Our works in this direction build on the graph-based framework to construct models that reflect the shape properties of the structure being modeled. We have used discriminative learning to exploit boundary- and symmetry-based representations for the construction of hierarchical models for shape detection, while for medical images we have developed methods for the end-to-end discriminative training of deformable contour models that combine low-level descriptors with contour-based organ boundary representations.

Regarding aspect (c) we have developed algorithms which implement top-down/bottom-up computation both in deterministic and stochastic optimization. The main idea is that 'bottom-up', image-based guidance is necessary for efficient detection, while 'top-down', object-based knowledge can disambiguate and help reliably interpret a given image; a combination of both modes of operation is necessary to combine accuracy with efficiency. In particular we have developed novel techniques for object detection that employ combinatorial optimization tools (A* and Branch-and-Bound) to tame the combinatorial complexity, achieving a best-case performance that is logarithmic in the number of pixels.

In the long run we aim at scaling up shape-based methods to 3D detection and pose estimation and large-scale object detection. One aspect which seems central to this is the development of appropriate mid-level representations. This is a problem that has received increased interest lately in the 2D case and is relatively mature, but in 3D it has been pursued primarily through ad-hoc schemes. We anticipate that questions pertaining to part sharing in 3D will be addressed most successfully by relying on explicit 3D representations. On the one hand depth sensors, such as Microsoft's Kinect, are now cheap enough to bring surface modeling and matching into the mainstream of computer vision - so these advances may be directly exploitable at test time for detection. On the other hand, even if we do not use depth information at test time, having 3D information can simplify the modeling task during training. In on-going work with collaborators we have started exploring combinations of such aspects, namely (i) the use of surface analysis tools to match surfaces from depth sensors (ii) using branch-and-bound for efficient inference in 3D space and (iii) groupwise-registration to build statistical 3D surface models. In the coming years we intend to pursue a tighter integration of these different directions for scalable 3D object recognition.

3.2. Machine Learning & Structured Prediction

The foundation of statistical inference is to learn a function that minimizes the expected loss of a prediction with respect to some unknown distribution

$$\mathcal{R}(f) = \int \ell(f, x, y) dP(x, y), \tag{2}$$

where $\ell(f, x, y)$ is a problem specific loss function that encodes a penalty for predicting $f(x)$ when the correct prediction is y. In our case, we consider x to be a medical image, and y to be some prediction, e.g. the segmentation of a tumor, or a kinematic model of the skeleton. The loss function, ℓ, is informed by the costs associated with making a specific misprediction. As a concrete example, if the true spatial extent of a tumor is encoded in y, $f(x)$ may make mistakes in classifying healthy tissue as a tumor, and mistakes in classifying diseased tissue as healthy. The loss function should encode the potential physiological damage resulting from erroneously targeting healthy tissue for irradiation, as well as the risk from missing a portion of the tumor.

A key problem is that the distribution P is unknown, and any algorithm that is to estimate f from labeled training examples must additionally make an implicit estimate of P. A central technology of empirical inference is to approximate $\mathcal{R}(f)$ with the empirical risk,

$$\mathcal{R}(f) \approx \widehat{\mathcal{R}}(f) = \frac{1}{n} \sum_{i=1}^{n} \ell(f, x_i, y_i), \tag{3}$$

which makes an implicit assumption that the training samples (x_i, y_i) are drawn i.i.d. from P. Direct minimization of $\widehat{\mathcal{R}}(f)$ leads to overfitting when the function class $f \in \mathcal{F}$ is too rich, and regularization is required:

$$\min_{f \in \mathcal{F}} \lambda \Omega(\|f\|) + \widehat{\mathcal{R}}(f), \tag{4}$$

where Ω is a monotonically increasing function that penalizes complex functions.

Equation Eq. 4 is very well studied in classical statistics for the case that the output, $y \in \mathcal{Y}$, is a binary or scalar prediction, but this is not the case in most medical imaging prediction tasks of interest. Instead, complex interdependencies in the output space leads to difficulties in modeling inference as a binary prediction problem. One may attempt to model e.g. tumor segmentation as a series of binary predictions at each voxel in a medical image, but this violates the i.i.d. sampling assumption implicit in Equation Eq. 3. Furthermore, we typically gain performance by appropriately modeling the inter-relationships between voxel predictions, e.g. by incorporating pairwise and higher order potentials that encode prior knowledge about the problem domain. It is in this context that we develop statistical methods appropriate to structured prediction in the medical imaging setting.

3.3. Self-Paced Learning with Missing Information

Many tasks in artificial intelligence are solved by building a model whose parameters encode the prior domain knowledge and the likelihood of the observed data. In order to use such models in practice, we need to estimate its parameters automatically using training data. The most prevalent paradigm of parameter estimation is supervised learning, which requires the collection of the inputs x_i and the desired outputs y_i. However, such an approach has two main disadvantages. First, obtaining the ground-truth annotation of high-level applications, such as a tight bounding box around all the objects present in an image, is often expensive. This prohibits the use of a large training dataset, which is essential for learning the existing complex models. Second, in many applications, particularly in the field of medical image analysis, obtaining the ground-truth annotation may not be feasible. For example, even the experts may disagree on the correct segmentation of a microscopical image due to the similarities between the appearance of the foreground and background.

In order to address the deficiencies of supervised learning, researchers have started to focus on the problem of parameter estimation with data that contains hidden variables. The hidden variables model the missing information in the annotations. Obtaining such data is practically more feasible: image-level labels ('contains car','does not contain person') instead of tight bounding boxes; partial segmentation of medical images. Formally, the parameters **w** of the model are learned by minimizing the following objective:

$$\min_{\mathbf{w} \in \mathcal{W}} R(\mathbf{w}) + \sum_{i=1}^{n} \Delta(y_i, y_i(\mathbf{w}), h_i(\mathbf{w})). \tag{5}$$

Here, $\mathcal{W}$ represents the space of all parameters, n is the number of training samples, $R(\cdot)$ is a regularization function, and $\Delta(\cdot)$ is a measure of the difference between the ground-truth output y_i and the predicted output and hidden variable pair $(y_i(\mathbf{w}), h_i(\mathbf{w}))$.

Previous attempts at minimizing the above objective function treat all the training samples equally. This is in stark contrast to how a child learns: first focus on easy samples ('learn to add two natural numbers') before moving on to more complex samples ('learn to add two complex numbers'). In our work, we capture this intuition using a novel, iterative algorithm called self-paced learning (SPL). At an iteration t, SPL minimizes the following objective function:

$$\min_{\mathbf{w}\in\mathcal{W},\mathbf{v}\in\{0,1\}^n} R(\mathbf{w}) + \sum_{i=1}^{n} v_i \Delta(y_i, y_i(\mathbf{w}), h_i(\mathbf{w})) - \mu_t \sum_{i=1}^{n} v_i. \tag{6}$$

Here, samples with $v_i = 0$ are discarded during the iteration t, since the corresponding loss is multiplied by 0. The term μ_t is a threshold that governs how many samples are discarded. It is annealed at each iteration, allowing the learner to estimate the parameters using more and more samples, until all samples are used. Our results already demonstrate that SPL estimates accurate parameters for various applications such as image classification, discriminative motif finding, handwritten digit recognition and semantic segmentation. We will investigate the use of SPL to estimate the parameters of the models of medical imaging applications, such as segmentation and registration, that are being developed in the GALEN team. The ability to handle missing information is extremely important in this domain due to the similarities between foreground and background appearances (which results in ambiguities in annotations). We will also develop methods that are capable of minimizing more general loss functions that depend on the (unknown) value of the hidden variables, that is,

$$\min_{\mathbf{w}\in\mathcal{W},\theta\in\Theta} R(\mathbf{w}) + \sum_{i=1}^{n} \sum_{h_i\in\mathcal{H}} \Pr(h_i | x_i, y_i; \theta) \Delta(y_i, h_i, y_i(\mathbf{w}), h_i(\mathbf{w})). \tag{7}$$

Here, θ is the parameter vector of the distribution of the hidden variables h_i given the input x_i and output y_i, and needs to be estimated together with the model parameters **w**. The use of a more general loss function will allow us to better exploit the freely available data with missing information. For example, consider the case where y_i is a binary indicator for the presence of a type of cell in a microscopical image, and h_i is a tight bounding box around the cell. While the loss function $\Delta(y_i, y_i(\mathbf{w}), h_i(\mathbf{w}))$ can be used to learn to classify an image as containing a particular cell or not, the more general loss function $\Delta(y_i, h_i, y_i(\mathbf{w}), h_i(\mathbf{w}))$ can be used to learn to detect the cell as well (since h_i models its location)

3.4. Discrete Biomedical Image Perception

A wide variety of tasks in medical image analysis can be formulated as discrete labeling problems. In very simple terms, a discrete optimization problem can be stated as follows: we are given a discrete set of variables $\mathcal{V}$, all of which are vertices in a graph $\mathcal{G}$. The edges of this graph (denoted by $\mathcal{E}$) encode the variables' relationships. We are also given as input a discrete set of labels $\mathcal{L}$. We must then assign one label from $\mathcal{L}$ to each variable in $\mathcal{V}$. However, each time we choose to assign a label, say, x_{p_1} to a variable p_1, we are forced to pay a price according to the so-called *singleton* potential function $g_p(x_p)$, while each time we choose to assign a pair of labels, say, x_{p_1} and x_{p_2} to two interrelated variables p_1 and p_2 (two nodes that are connected by an edge in the graph $\mathcal{G}$), we are also forced to pay another price, which is now determined by the so called *pairwise* potential function $f_{p_1p_2}(x_{p_1}, x_{p_2})$. Both the singleton and pairwise potential functions are problem specific and are thus assumed to be provided as input.

Our goal is then to choose a labeling which will allow us to pay the smallest total price. In other words, based on what we have mentioned above, we want to choose a labeling that minimizes the sum of all the MRF potentials, or equivalently the MRF energy. This amounts to solving the following optimization problem:

$$\arg\min_{\{x_p\}} \mathcal{P}(g,f) = \sum_{p\in\mathcal{V}} g_p(x_p) + \sum_{(p_1,p_2)\in\mathcal{E}} f_{p_1p_2}(x_{p_1},x_{p_2}). \tag{8}$$

The use of such a model can describe a number of challenging problems in medical image analysis. However these simplistic models can only account for simple interactions between variables, a rather constrained scenario for high-level medical imaging perception tasks. One can augment the expression power of this model through higher order interactions between variables, or a number of cliques $\{C_i, i\in[1,n]=\{\{p_{i^1},\cdots,p_{i^{|C_i|}}\}\}$ of order $|C_i|$ that will augment the definition of $\mathcal{V}$ and will introduce hyper-vertices:

$$\arg\min_{\{x_p\}} \mathcal{P}(g,f) = \sum_{p\in\mathcal{V}} g_p(x_p) + \sum_{(p_1,p_2)\in\mathcal{E}} f_{p_1p_2}(x_{p_1},x_{p_2}) + \sum_{C_i\in\mathcal{E}} f_{p_1\cdots p_n}(x_{p_{i^1}},\cdots,p_{x_{i^{|C_i|}}}). \tag{9}$$

where $f_{p_1\cdots p_n}$ is the price to pay for associating the labels $(x_{p_{i^1}},\cdots,p_{x_{i^{|C_i|}}})$ to the nodes $(p_1\cdots p_{i^{|C_i|}})$. Parameter inference, addressed by minimizing the problem above, is the most critical aspect in computational medicine and efficient optimization algorithms are to be evaluated both in terms of computational complexity as well as of inference performance. State of the art methods include deterministic and non-deterministic annealing, genetic algorithms, max-flow/min-cut techniques and relaxation. These methods offer certain strengths while exhibiting certain limitations, mostly related to the amount of interactions which can be tolerated among neighborhood nodes. In the area of medical imaging where domain knowledge is quite strong, one would expect that such interactions should be enforced at the largest scale possible.

4. Application Domains

4.1. Breast tomosynthesis

Participants: Emilie Chouzenoux, Jean-Christophe Pesquet, Maissa Sghaier (collaboration G. Palma, GE Healthcare)

Breast cancer is the most frequently diagnosed cancer for women. Mammography is the most used imagery tool for detecting and diagnosing this type of cancer. Since it consists of a 2D projection method, this technique is sensitive to geometrical limitations such as the superimposition of tissues which may reduce the visibility of lesions or make even appear false structures which are interpreted by radiologists as suspicious signs. Digital breast tomosynthesis allows these limitations to be circumvented. This technique is grounded on the acquisition of a set of projections with a limited angle view. Then, a 3D estimation of the sensed object is performed from this set of projections, so reducing the overlap of structures and improving the visibility and detectability of lesions possibly present in the breast. The objective of our work is to develop a high quality reconstruction methodology where the full pipeline of data processing will be modeled.

4.2. Inference of gene regulatory networks

Participants: Jean-Christophe Pesquet (collaboration A. Pirayre and L. Duval, IFPEN)

The discovery of novel gene regulatory processes improves the understanding of cell phenotypic responses to external stimuli for many biological applications, such as medicine, environment or biotechnologies. To this purpose, transcriptomic data are generated and analyzed from DNA microarrays or more recently RNAseq experiments. They consist in genetic expression level sequences obtained for all genes of a studied organism placed in dierent living conditions. From these data, gene regulation mechanisms can be recovered by revealing topological links encoded in graphs. In regulatory graphs, nodes correspond to genes. A link between two nodes is identified if a regulation relationship exists between the two corresponding genes. In our work, we propose to address this network inference problem with recently developed techniques pertaining to graph optimization. Given all the pairwise gene regulation information available, we propose to determine the presence of edges in the considered GRN by adopting an energy optimization formulation integrating additional constraints. Either biological (information about gene interactions) or structural (information about node connectivity) a priori are considered to restrict the space of possible solutions. Different priors lead to different properties of the global cost function, for which various optimization strategies, either discrete and continuous, can be applied.

4.3. Lung Tumor Detection and Characterization

Participants: Evgenios Kornaropoulos, Evangelia Zacharaki, Nikos Paragios

The use of Diffusion Weighted MR Imaging (DWI) is investigated as an alternative tool to radiologists for tumor detection, tumor characterization, distinguishing tumor tissue from non-tumor tissue, and monitoring and predicting treatment response. In collaboration with Hôpitaux Universitaires Henri-Mondor in Paris, France and Chang Gung Memorial Hospital – Linkou in Taipei, Taiwan we investigate the use of modelbased methods of 3D image registration, clustering and segmentation towards the development of a framework for automatic interpretation of images, and in particular extraction of meaningful biomarkers in aggressive lymphomas.

4.4. Protein function prediction

Participants: Evangelia Zacharaki, Nikos Paragios (in collaboration with D. Vlachakis, University of Patras, Greece)

The massive expansion of the worldwide Protein Data Bank (PDB) provides new opportunities for computational approaches which can learn from available data and extrapolate the knowledge into new coming instances. The aim of our work was to exploit experimentally acquired structural information of enzymes through machine learning techniques in order to produce models that predict enzymatic function.

4.5. Imaging biomarkers for chronic lung diseases

Participants: Guillaume Chassagnon, Evangelia Zacharaki, Maria Vakalopoulou, Nikos Paragios

Diagnosis and staging of chronic lung diseases is a major challenge for both patient care and approval of new treatments. Among imaging techniques, computed tomography (CT) is the gold standard for in vivo morphological assessment of lung parenchyma currently offering the highest spatial resolution in chronic lung diseases. Although CT is widely used its optimal use in clinical practice and as an endpoint in clinical trials remains controversial. Our goal is to develop quantitative imaging biomarkers that allow (i) severity assessment (based on the correlation to functional and clinical data) and (ii) monitoring the disease progression. In the current analysis we focus on scleroderma and cystic fibrosis as models for restrictive and obstructive lung disease, respectively. Two different approaches are investigated: disease assessment by histogram or texture analysis and assessment of the regional lung elasticity through deformable registration. This work is in collaboration with the Department of Radiology, Cochin Hospital, Paris.

4.6. Co-segmentation and Co-registration of Subcortical Brain Structures

Participants: Enzo Ferrante, Nikos Paragios, Iasonas Kokkinos

New algorithms to perform co-segmentation and co-registration of subcortical brain structures on MRI images were investigated in collaboration with Ecole Polytechnique de Montreal and the Sainte-Justine Hospital Research Center from Montreal. Brain subcortical structures are involved in different neurodegenerative and neuropsychiatric disorders, including schizophrenia, Alzheimers disease, attention deficit, and subtypes of epilepsy. Segmenting these parts of the brain enables a physician to extract indicators, facilitating their quantitative analysis and characterization. We are investigating how estimated maps of semantic labels (obtained using machine learning techniques) can be used as a surrogate for unlabelled data. We are exploring how to combine them with multi-population deformable registration to improve both alignment and segmentation of these challenging brain structures.

4.7. Restoration of old video archives

Participants: Emilie Chouzenoux, Jean-Christophe Pesquet (collaboration F. Abboud, WITBE, J.-H. Chenot and L. Laborelli, INA)

The last century has witnessed an explosion in the amount of video data stored with holders such as the National Audiovisual Institute whose mission is to preserve and promote the content of French broadcast programs. the cultural impact of these records, their value is increased due to commercial reexploitation through recent visual media. However, the perceived quality of the old data fails to satisfy the current public demand. The purpose of our work is to propose new methods for restoring video sequences supplied from television archive documents, using modern optimization techniques with proven convergence properties.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

The work on dense registration of faces [22] was selected as demo at the IEEE Conference on Computer Vision and Pattern Recognition.

The work [26] received the best poster award at the BASP workshop 2017.

5.1.2. Others

Emilie Chouzenoux received an ANR JCJC grant, for her project MajIC:Majorization-Minimization algorithms for Image Computing.

Evangelia Zacharaki has defended her 'Habilitation à Diriger des Recherches' [3].

Emilie Chouzenoux has defended her 'Habilitation à Diriger des Recherches' [1].

6. New Software and Platforms

6.1. DISD

Dense Image and Surface Descriptors

FUNCTIONAL DESCRIPTION: Scale-Invariant Descriptor, Scale-Invariant Heat Kernel Signatures DISD implements the SID, SI-HKS and ISC descriptors. SID (Scale-Invariant Descriptor) is a densely computable, scale- and rotation- invariant descriptor. We use a log-polar grid around every point to turn rotation/scalings into translation, and then use the Fourier Transform Modulus (FTM) to achieve invariance. SI-HKS (Scale-Invariant Heat Kernel Signatures) extract scale-invariant shape signatures by exploiting the fact that surface scaling amounts to multiplication and scaling of a properly sampled HKS descriptor. We apply the FTM trick on HKS to achieve invariance to scale changes. ISC (Intrinsic Shape Context) constructs a net-like grid around every surface point by shooting outwards and tracking geodesics. This allows us to build a meta-descriptor on top of HKS/SI-HKS that takes neighborhood into account, while being invariant to surface isometries.

- Participants: Eduard Trulls and Iasonas Kokkinos
- Contact: Iasonas Kokkinos
- URL: http://vision.mas.ecp.fr/Personnel/iasonas/descriptors.html

6.2. DPMS

FUNCTIONAL DESCRIPTION: Dpms implements branch-and-bound object detection, cutting down the complexity of detection from linear in the number of pixels to logarithmic.

- Participant: Iasonas Kokkinos
- Contact: Iasonas Kokkinos

6.3. DROP

KEYWORDS: Health - Merging - Registration of 2D and 3D multimodal images - Medical imaging
FUNCTIONAL DESCRIPTION: Drop is a software programme that registers images originating from one or more modes by quickly and efficiently calculating a non-rigid / deformable field of deformation. Drop is a new, quick and effective registration tool based on new algorithms that do not require a cost function derivative.

- Partner: Centrale Paris
- Contact: Nikolaos Paragyios
- URL: http://campar.in.tum.de/Main/Drop

6.4. FastPD

KEYWORD: Medical imaging
FUNCTIONAL DESCRIPTION: FastPD is an optimization platform in C++ for the computer vision and medical imaging community.

- Contact: Nikolaos Paragyios
- URL: http://www.csd.uoc.gr/~komod/FastPD/

6.5. GraPeS

Grammar Parser for Shapes
FUNCTIONAL DESCRIPTION: It is a software for parsing facade images using shape grammars. Grapes implement a parsing methods based on Reinforcement Learning principles. It optimizes simultaneously the topology of the parse tree as well as the associated parameters. GraPeS comes along with predefined shape grammars as XML files and defines three kinds of rewards. However, it also offers the possibility to create new grammars and to provide custom rewards in text files, widening the scope of potential applications. The name of the software comes from the aspect of the parse tree of the binary split grammars involved in the process.

RELEASE FUNCTIONAL DESCRIPTION: Supports jpg and gif file formats.

- Participant: Iasonas Kokkinos
- Contact: Iasonas Kokkinos

6.6. HOAP-SVM

High-Order Average Precision SVM
SCIENTIFIC DESCRIPTION: We consider the problem of using high-order information (for example, persons in the same image tend to perform the same action) to improve the accuracy of ranking (specifically, average precision). We develop two learning frameworks. The high-order binary SVM (HOB-SVM) optimizes a convex upper bound of the surrogate 0-1 loss function. The high-order average precision SVM (HOAP-SVM) optimizes a difference-of-convex upper bound on the average precision loss function.

Authors of the research paper: Puneet K. Dokania, A. Behl, C. V. Jawahar and M. Pawan Kumar

FUNCTIONAL DESCRIPTION: The software provides a convenient API for learning to rank with high-order information. The samples are ranked according to a score that is proportional to the difference of max-marginals of the positive and the negative class. The parameters of the score function are computed by minimizing an upper bound on the average precision loss. The software also provides an instantiation of the API for ranking samples according to their relevance to an action, using the poselet features. The following learning algorithms are included in the API:

(1) Multiclass-SVM (2) AP-SVM (3) High Order Binary SVM (HOB-SVM) (4) High Order AP-SVM (HOAP-SVM) (5) M4 Learning (unpublished work)

The API is developed in C/C++ by Puneet K. Dokania.

- Participants: Pawan Kumar and Puneet Dokania
- Contact: Puneet Dokania
- URL: http://puneetkdokania.github.io/projects/ranking-highorder/ranking-highorder.html

6.7. LBSD

Learning-Based Symmetry Detection
FUNCTIONAL DESCRIPTION: LBSD implements the learning-based approach to symmetry detection. It includes the code for running a detector, alongside with the ground-truth symmetry annotations that we have introduced for the Berkeley Segmentation Dataset (BSD) benchmark.

- Participant: Stavros Tsogkas
- Contact: Stavros Tsogkas
- URL: https://github.com/tsogkas/oid_1.0

6.8. mrf-registration

KEYWORDS: Health - Medical imaging
FUNCTIONAL DESCRIPTION: Deformable image and volume registration, is a deformable registration platform in C++ for the medical imaging community. This is the first publicly available platform which contains most of the existing metrics to perform registration under the same concept. The platform is used for clinical research from approximately 3,000 users worldwide.

RELEASE FUNCTIONAL DESCRIPTION: Bugfix in image resampling Resampling of binary mask is now w.r.t. to target image Added adjustable sigma for Gaussian image pyramid Added level dependent scaling of maximum displacement when linkMax is disabled Changed approximation method for computation of the inverse displacment field (less memory demanding) Bugfix in grid and quiver visualization Added support for compressed MHD Bugfix in 3D thin-plate splines in landmark-based registration

- Participant: Nikolaos Paragyios
- Contact: Nikolaos Paragyios
- URL: http://www.mrf-registration.net/

6.9. TeXMeG

FUNCTIONAL DESCRIPTION: Texture, modulation, generative models, segmentation, TeXMeG is a front-end for texture analysis and edge detection platform in Matlab that relies on Gabor filtering and image demodulation. Includes frequency- and time- based definition of Gabor- and other Quadrature-pair filterbanks, demodulation with the Regularized Energy Separation Algorithm and Texture/Edge/Smooth classification based on MDL criterion.

- Participant: Iasonas Kokkinos
- Contact: Iasonas Kokkinos
- URL: http://cvsp.cs.ntua.gr/software/texture/

6.10. Platforms

6.10.1. The Proximity Operator Repository

Participants: Emilie Chouzenoux and Jean-Christophe Pesquet (in collaboration with Giovanni Chierchia, Univ. Paris Est, and Patrick Combettes, North Carolina State University).
link: http://proximity-operator.net/
Proximity operators have become increasingly important tools as basic building blocks of proximal splitting algorithms, a class of algorithms that decompose complex composite convex optimization methods into simple steps involving one of the functions present in the model. This website provides formulas for efficiently computing the proximity operator of various functions, along with the associated codes.

7. New Results

7.1. Graph Based Slice-to-Volume Deformable Registration

Participants: Enzo Ferrante, Nikos Paragios
Deformable image registration is a fundamental problem in computer vision and medical image computing. In this contribution [9], we investigate the use of graphical models in the context of a particular type of image registration problem, known as slice-to-volume registration, while we introduced the first comprehensive survey [10] of the literature about slice-to-volume registration, presenting a categorical study of the algorithms according to an ad-hoc taxonomy and analyzing advantages and disadvantages of every category. We introduce a scalable, modular and flexible formulation that can accommodate low-rank and high order terms, that simultaneously selects the plane and estimates the in-plane deformation through a single shot optimization approach. The proposed framework is instantiated into different variants seeking either a compromise between computational efficiency (soft plane selection constraints and approximate definition of the data similarity terms through pair-wise components) or exact definition of the data terms and the constraints on the plane selection. Simulated and real-data in the context of ultrasound and magnetic resonance registration (where both framework instantiations as well as different optimization strategies are considered) demonstrate the potentials of our method.

7.2. Deformable Registration Through Learning of Context-Specific Metric Aggregation

Participants: Enzo Ferrante, Rafael Marini, Punnet K. Dokania, Nikos Paragios
We propose a novel weakly supervised discriminative algorithm [21] for learning context specific registration metrics as a linear combination of conventional similarity measures. Conventional metrics have been extensively used over the past two decades and therefore both their strengths and limitations are known. The challenge is to find the optimal relative weighting (or parameters) of different metrics forming the similarity measure of the registration algorithm. Hand-tuning these parameters would result in sub optimal solutions and quickly become infeasible as the number of metrics increases. Furthermore, such hand-crafted combination can only happen at global scale (entire volume) and therefore will not be able to account for the different tissue properties. We propose a learning algorithm for estimating these parameters locally, conditioned to the data semantic classes. The objective function of our formulation is a special case of non-convex function, difference of convex function, which we optimize using the concave convex procedure. As a proof of concept, we show the impact of our approach on three challenging datasets for different anatomical structures and modalities.

7.3. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology

Participants: Roger Sun, Evangelia I. Zacharaki, Nikos Paragios (in collaboration with Gustave Roussy and Paris Sud University)

Computational medical imaging (also known as radiomics) is a promising and rapidly growing discipline that consists in the analysis of high-dimensional data extracted from medical imaging, to further describe tumour phenotypes. The end goal of radiomics is to determine imaging biomarkers as decision support tools for clinical practice and to facilitate better understanding of cancer biology, allowing the assessment of the changes throughout the evolution of the disease and the therapeutic sequence. We have reviewed [12], [17] the critical issues necessary for proper development of radiomics as a biomarker and for its implementation in clinical practice.

7.4. Multi-atlas segmentation in medical imagery

Participants: Stavros Alchatzidis, Evangelia I. Zacharaki, Nikos Paragios (in collaboration with University of Pennsylvania)

Multi-atlas segmentation has emerged in recent years as a simple yet powerful approach in medical image segmentation. It commonly comprises two steps: (1) a series of pairwise registrations that establish correspondences between a query image and a number of atlases, and (2) the fusion of the available segmentation hypotheses towards labeling objects of interest. In [5], we introduce a novel approach that solves simultaneously for the underlying segmentation labels and the multi-atlas registration. We propose a pairwise Markov Random Field approach,where registration and segmentation nodes are coupled towards simultaneously recovering all atlas deformations and labeling the query image.

7.5. Protein function prediction

Participants: Evangelia I. Zacharaki, Nikos Paragios (in collaboration with University of Patras)

The massive expansion of the worldwide Protein Data Bank (PDB) provides new opportunities for computational approaches which can learn from available data and extrapolate the knowledge into new coming instances. The aim of our work in [6], [18] was to exploit experimentally acquired structural information of enzymes through machine learning techniques in order to produce models that predict enzymatic function.

7.6. Deformable group-wise registration using a physiological model: Application to diffusion-weighted MRI

Participants: Evgenios Kornaropoulos, Evangelia I. Zacharaki, Nikos Paragios (in collaboration with Centre Hospitalier Universitaire Henri-Mondor and Chang Gung Memorial Hospital)

In this contribution [2] we develop a novel group-wise deformable registration method for motion correction in Diffusion-Weighted MRI towards computing a more accurate Apparent Diffusion Coefficient parametric map (ADC map). Calculation of the ADC has been performed without motion correction in the previous studies. It is reported though that ADC is a parameter susceptible to artifacts, the most frequent of all being patient's motion and breathing, resulting in misregistration of the images obtained with different b-values. Being group-wise designed, the image registration method we propose has no need of choosing a reference template while in the same time it is computationally efficient. We aim at finding the optimal deformation fields of the diffusion-weighted (DW) images using a temporal constraint, related to the diffusion process, as well as a smoothness penalty on the deformations. To this end, we address the deformation fields estimation problem with an Markov Random Fields formulation, in which the latent variables are the deformations (B-spline polynomials) of the images. The latent variables are connected with the observations towards ensuring meaningful temporal correspondence among the DW images. They are also inter-connected in order to decrease the cost of pairwise comparisons between individual images. Linear programming and duality are used to determine the optimal solution of the problem. Finally, as an image similarity criterion in the MRF framework, we used a metric that was based on a physiological model describing the image acquisition process. Quantitative evaluation of the method was performed, in which it was compared against two state-of-the-art methods that use other modelling criteria, It outperformed both of them, while the ADC map derived by our method appeared to preserve structure, that was not observable by the other methods.

7.7. Variational Bayesian Approach for Image Restoration

Participants: Emilie Chouzenoux and Jean-Christophe Pesquet (in collaboration with Y. Marnissi, SAFRAN TECH and Y. Zheng, IBM Research China)

In the work [13], a methodology is investigated for signal recovery in the presence of non-Gaussian noise. In contrast with regularized minimization approaches often adopted in the literature, in our algorithm the regularization parameter is reliably estimated from the observations. As the posterior density of the unknown parameters is analytically intractable, the estimation problem is derived in a variational Bayesian framework where the goal is to provide a good approximation to the posterior distribution in order to compute posterior mean estimates. Moreover, a majorization technique is employed to circumvent the difficulties raised by the intricate forms of the non-Gaussian likelihood and of the prior density. We demonstrate the potential of the proposed approach through comparisons with state-of-the-art techniques that are specifically tailored to signal recovery in the presence of mixed Poisson-Gaussian noise. Results show that the proposed approach is efficient and achieves performance comparable with other methods where the regularization parameter is manually tuned from an available ground truth.

7.8. Non-Modular Loss Functions

Participant: Jiaqian Yu

Defining Non-modular loss functions and their optimization procedure present an interesting direction for many classes of problems. Jiaqian Yu has completed her PhD Thesis on Non-Modular Loss Functions this year. The PhD Thesis has included several yet unpublished results regarding approximate losses for Jaccard index and DICE coefficients commonly used in evaluating segmentation algorithms.

7.9. Graph Structure Discovery

Participant: Eugene Belilovsky

Discovering the interaction structure amongst variables, particularly from few observations, has important implications in many fields including neuroimaging, genetics and finance. Eugene Belilovsky in collaboration with Gael Varoquaux (Inria Parietal), Kyle Kastner (University of Montreal) and Matthew Blaschko has published a new approach for graph structure discovery in high dimensional gaussian markov random fields. The work has been presented in [19].

7.10. Structured and Efficient Convolutional Networks

Participant: Eugene Belilovsky

Convolutional Neural Networks have revolutionized the computer vision field. Yet, they are not well understood and do not well leverage basic geometric structures known by the computer vision community. In recent work in collaboration with the Ecole Normale Superier and the École des Ponts ParisTech we have tried to address some of these issues. We use as a starting point the recently introduced Scattering Transform and show that we can use this to build Convolutional Networks that are more interpertable and can generalize faster in the few sample regime. This work has been presented in [25].

7.11. Stochastic Majorize-Minimize Subspace Algorithm

Participants: Emilie Chouzenoux and Jean-Christophe Pesquet

Stochastic optimization plays an important role in solving many problems encountered in machine learning or adaptive processing. In this context, the second-order statistics of the data are often unknown a priori or their direct computation is too intensive, and they have to be estimated on-line from the related signals. In the context of batch optimization of an objective function being the sum of a data fidelity term and a penalization (e.g. a sparsity promoting function), Majorize-Minimize (MM) subspace methods have recently attracted much interest since they are fast, highly flexible and effective in ensuring convergence. The goal of the work [8] is to show how these methods can be successfully extended to the case when the cost function is replaced by a sequence of stochastic approximations of it. Simulation results illustrate the good practical performance of the proposed MM Memory Gradient (3MG) algorithm when applied to 2D filter identification

7.12. Deconvolution and Deinterlacing of Video Sequences

Participants: Emilie Chouzenoux and Jean-Christophe Pesquet (in collaboration with F. Abboud, PhD student, J.-H. Chenot and L. Laborelli, research engineers, Institut National de l'Audiovisuel)

Optimization methods play a central role in the solution of a wide array of problems encountered in various application fields, such as signal and image processing. Especially when the problems are highly dimensional, proximal methods have shown their efficiency through their capability to deal with composite, possibly non smooth objective functions. The cornerstone of these approaches is the proximity operator, which has become a quite popular tool in optimization. In the work [31], we propose new dual forward-backward formulations for computing the proximity operator of a sum of convex functions involving linear operators. The proposed algorithms are accelerated thanks to the introduction of a block coordinate strategy combined with a preconditioning technique. Numerical simulations emphasize the good performance of our approach for the problem of jointly deconvoluting and deinterlacing video sequences.

7.13. PALMA, an improved algorithm for DOSY signal processing

Participants: Emilie Chouzenoux (in collaboration with M.-A. Delsuc, IGBMC, Strasbourg, and A. Cherni, PhD student, Univ. Strasbourg)

NMR is a tool of choice for the measure of diffusion coefficients of species in solution. The DOSY experiment, a 2D implementation of this measure, has proven to be particularly useful for the study of complex mixtures, molecular interactions, polymers, etc. However, DOSY data analysis requires to resort to inverse Laplace transform, in particular for polydisperse samples. This is a known difficult numerical task, for which we present here a novel approach. A new algorithm based on a splitting scheme and on the use of proximity operators is introduced in [7]. Used in conjunction with a Maximum Entropy and ℓ_1 hybrid regularisation, this algorithm converges rapidly and produces results robust against experimental noise. This method has been called PALMA. It is able to reproduce faithfully monodisperse as well as polydisperse systems, and numerous simulated and experimental examples are presented. It has been implemented on the server http://palma.labo.igbmc.fr where users can have their datasets processed automatically.

7.14. Proximal Approaches for Solving Matrix Optimization Problems

Participants: Emilie Chouzenoux and Jean-Christophe Pesquet (in collaboration with A. Benfenati, Univ. Paris Est)

In recent years, there has been a growing interest in problems where the underlining mathematical model involves the minimization in a matrix space of a Bregman divergence function coupled with a regularization term. We consider a general framework where the regularization term is decoupled in two parts, one acting only on the eigenvalues of the matrix and the other on the whole matrix. We propose in [26], [32] a new minimization approach to address problem of this type, by providing a list of proximity operators allowing us to consider various choices for the fit–to–data functional and for the regularization term. The numerical experience show that this approach gives better results in term of computational time with respect to some state of the arts algorithms.

7.15. Fast Algorithm for Least-Squares Regression with GMRF Prior

Participants: Emilie Chouzenoux and Jean-Christophe Pesquet (in collaboration with J.Y. Tourneret, IRIT, Toulouse, and Q. Wei, Duke Univ.)

The paper [29] presents a fast approach for penalized least squares (LS) regression problems using a 2D Gaussian Markov random field (GMRF) prior. More precisely, the computation of the proximity operator of the LS criterion regularized by different GMRF potentials is formulated as solving a Sylvester-like matrix equation. By exploiting the structural properties of GMRFs, this matrix equation is solved column-wise in an analytical way. The proposed algorithm can be embedded into a wide range of proximal algorithms to solve LS regression problems including a convex penalty. Experiments carried out in the case of a constrained

LS regression problem arising in a multichannel image processing application, provide evidence that an alternating direction method of multipliers performs quite efficiently in this context.

7.16. Optimization Approach for Deep Neural Network Training

Participants: Emilie Chouzenoux, Jean-Christophe Pesquet, Vyacheslav Dudar (in collaboration with G. Chierchia, Univ. Paris Est and V. Semenov, Univ. of Kiev)

In paper [28], we develop a novel second-order method for training feed-forward neural nets. At each iteration, we construct a quadratic approximation to the cost function in a low-dimensional subspace. We minimize this approximation inside a trust region through a two-stage procedure: first inside the embedded positive curvature subspace, followed by a gradient descent step. This approach leads to a fast objective function decay, prevents convergence to saddle points, and alleviates the need for manually tuning parameters. We show the good performance of the proposed algorithm on benchmark datasets.

7.17. Auxiliary Variable Method for MCMC Algorithms in High Dimension

Participants: Emilie Chouzenoux and Jean-Christophe Pesquet (in collaboration with Y. Marnissi, SAFRAN TECH and A. Benazza-Benhayia, SUP'COM, COSIM, Tunis)

When the parameter space is high dimensional, the performance of stochastic sampling algorithms is very sensitive to existing dependencies between parameters. For instance, this problem arises when one aims to sample from a high dimensional Gaussian distribution whose covariance matrix does not present a simple structure. Then, one often resorts to sampling algorithms based on a perturbation-optimization technique that requires to minimize a cost function using an iterative algorithm. This makes the sampling process time consuming, especially when used within a Gibbs sampler. Another challenge is the design of Metropolis-Hastings proposals that make use of information about the local geometry of the target density in order to speed up the convergence and improve mixing properties in the parameter space, while being not too computationally expensive. These two contexts are mainly related to the presence of two heterogeneous sources of dependencies stemming either from the prior or the likelihood in the sense that the related covariances matrices cannot be diagonalized in the same basis. In paper [34], we are interested in inverse problems where either the data fidelity term or the prior distribution is Gaussian or driven from a hierarchical Gaussian model. We propose to add auxiliary variables to the model in order to dissociate the two sources of dependencies. In the new augmented space, only one source of correlation remains directly related to the target parameters, the other sources of correlations being captured by the auxiliary variables. Experiments conducted on two image restoration problems show the good performance of the proposed strategy.

7.18. Block Coordinate Approach for Sparse Logistic Regression

Participants: Emilie Chouzenoux and Jean-Christophe Pesquet (in collaboration with G. Chierchia, Univ. Paris Est, L. M. Briceno-Arias, CMM - Univ. Chile, and A. Cherni, PhD student, Univ. Strasbourg)

We propose in [20], [33] stochastic optimization algorithms for logistic regression based on a randomized version of Douglas–Rachford splitting method. Our approach sweeps the training set by randomly selecting a mini-batch of data at each iteration, and it performs the update step by leveraging the proximity operator of the logistic loss, for which a closed-form expression is derived. Experiments carried out on standard datasets compare the efficiency of our algorithm to stochastic gradient-like methods.

7.19. An Alternating Proximal Approach for Blind Video Deconvolution

Participants: Emilie Chouzenoux and Jean-Christophe Pesquet (in collaboration with Feriel Abboud, WITBE, Jean-Hugues Chenot, Louis Laborelli, INA)

Blurring occurs frequently in video sequences captured by consumer devices, as a result of various factors such as lens aberrations, defocus, relative camera scene motion, and camera shake. When it comes to the contents of archive documents such as old films and television shows, the degradations are even more serious due to several physical phenomena happening during the sensing, transmission, recording, and storing processes. We propose in [31] a versatile formulation of blind video deconvolution problems that seeks to estimate both the sharp unknown video sequence and the underlying blur kernel from an observed video. This inverse problem is severely ill-posed, and an appropriate solution can be obtained by modeling it as a nonconvex minimization problem. We provide a novel iterative algorithm to solve it, grounded on the use of recent advances in convex and nonconvex optimization techniques, and having the ability of including numerous well-known regularization strategies

7.20. BRANE Clust: Cluster-Assisted Gene Regulatory Network Inference Refinement

Participants: Jean-Christophe Pesquet (in collaboration with Aurélie Pirayre, IFP Energies nouvelles, Camille Couprie, Facebook Research, Laurent Duval, IFP Energies nouvelles)

Discovering meaningful gene interactions is crucial for the identification of novel regulatory processes in cells. Building accurately the related graphs remains challenging due to the large number of possible solutions from available data. Nonetheless, enforcing a priori on the graph structure, such as modularity, may reduce network indeterminacy issues. BRANE Clust (Biologically-Related A priori Network Enhancement with Clustering) refines gene regulatory network (GRN) inference thanks to cluster information. It works as a post-processing tool for inference methods (i.e. CLR, GENIE3). In BRANE Clust, the clustering is based on the inversion of a system of linear equations involving a graph-Laplacian matrix promoting a modular structure. Our approach [16] is validated on DREAM4 and DREAM5 datasets with objective measures, showing significant comparative improvements. We provide additional insights on the discovery of novel regulatory or co-expressed links in the inferred Escherichia coli network evaluated using the STRING database. The comparative pertinence of clustering is discussed computationally (SIMoNe, WGCNA, X-means) and biologically (RegulonDB).

7.21. Proximity Operators of Discrete Information Divergences

Participants: Jean-Christophe Pesquet (in collaboration with Mireille El Gheche, EPFL, Giovanni Chierchia, ESIEE Paris)

Information divergences allow one to assess how close two distributions are from each other. Among the large panel of available measures, a special attention has been paid to convex ϕ-divergences, such as Kullback-Leibler, Jeffreys-Kullback, Hellinger, Chi-Square, Renyi, and I_α divergences. While ϕ-divergences have been extensively studied in convex analysis, their use in optimization problems often remains challenging. In this regard, one of the main shortcomings of existing methods is that the minimization of ϕ-divergences is usually performed with respect to one of their arguments, possibly within alternating optimization techniques. In this paper, we overcome this limitation by deriving new closed-form expressions for the proximity operator of such two-variable functions. This makes it possible to employ standard proximal methods for efficiently solving a wide range of convex optimization problems involving ϕ-divergences. In addition, we show that these proximity operators are useful to compute the epigraphical projection of several functions of practical interest. The proposed proximal tools are numerically validated in the context of optimal query execution within database management systems, where the problem of selectivity estimation plays a central role. Experiments are carried out on small to large scale scenarios.

7.22. Stochastic Quasi-Fejér Block-Coordinate Fixed Point Iterations With Random Sweeping: Mean-Square and Linear Convergence

Participants: Jean-Christophe Pesquet (in collaboration with Patrick L. Combettes, North Caroline State University)

In one of our previous works, we investigated the almost sure weak convergence of a block-coordinate fixed point algorithm and discussed its application to nonlinear analysis and optimization. This algorithm features random sweeping rules to select arbitrarily the blocks of variables that are activated over the course of the iterations and it allows for stochastic errors in the evaluation of the operators. The present paper establishes results on the mean-square and linear convergence of the iterates. Applications to monotone operator splitting and proximal optimization algorithms are presented.

7.23. Human Joint Angle Estimation and Gesture Recognition for Assistive Robotic Vision

Participants: Riza Alp Guler, Siddhartha Chandra, Iasonas Kokkinos (in collaboration with National Technical University of Athens)

In this work, we explore new directions for automatic human gesture recognition and human joint angle estimation as applied for human-robot interaction in the context of an actual challenging task of assistive living for real-life elderly subjects. Our contributions include state-of-the-art approaches for both low- and mid-level vision, as well as for higher level action and gesture recognition. The first direction investigates a deep learning based framework for the challenging task of human joint angle estimation on noisy real world RGB-D images. The second direction includes the employment of dense trajectory features for online processing of videos for automatic gesture recognition with real-time performance. Our approaches are evaluated both qualitative and quantitatively on a newly acquired dataset that is constructed on a challenging real-life scenario on assistive living for elderly subjects.

7.24. Fast, Exact and Multi-Scale Inference for Semantic Image Segmentation with Deep Gaussian CRFs

Participants: Siddhartha Chandra, Iasonas Kokkinos

In this work we propose a structured prediction technique that combines the virtues of Gaussian Conditional Random Fields (G-CRF) with Deep Learning: (a) our structured prediction task has a unique global optimum that is obtained exactly from the solution of a linear system (b) the gradients of our model parameters are analytically computed using closed form expressions, in contrast to the memory-demanding contemporary deep structured prediction approaches that rely on back-propagation-through-time, (c) our pairwise terms do not have to be simple hand-crafted expressions, as in the line of works building on the DenseCRF, but can rather be ‘discovered’ from data through deep architectures, and (d) out system can trained in an end-to-end manner. Building on standard tools from numerical analysis we develop very efficient algorithms for inference and learning, as well as a customized technique adapted to the semantic segmentation task. This efficiency allows us to explore more sophisticated architectures for structured prediction in deep learning: we introduce multi-resolution architectures to couple information across scales in a joint optimization framework, yielding systematic improvements. We demonstrate the utility of our approach on the challenging VOC PASCAL 2012 image segmentation benchmark, showing substantial improvements over strong baselines.

7.25. Dense and Low-Rank Gaussian CRFs Using Deep Embeddings

Participants: Siddhartha Chandra, Iasonas Kokkinos

In this work we introduce a structured prediction model that endows the Deep Gaussian Conditional Random Field (G-CRF) with a densely connected graph structure. We keep memory and computational complexity under control by expressing the pairwise interactions as inner products of low-dimensional, learnable embeddings. The G-CRF system matrix is therefore low-rank, allowing us to solve the resulting system in a few milliseconds on the GPU by using conjugate gradient. As in G-CRF, inference is exact, the unary and pairwise terms are jointly trained end-to-end by using analytic expressions for the gradients, while we also develop even faster, Potts-type variants of our embeddings. We show that the learned embeddings capture pixel-to-pixel affinities in a task-specific manner, while our approach achieves state of the art results on three challenging benchmarks, namely semantic segmentation, human part segmentation, and saliency estimation. This work was published in [30].

7.26. DenseReg: Fully Convolutional Dense Shape Regression In-the-Wild

Participants: Riza Alp Guler, Iasonas Kokkinos (in collaboration with Imperial College London)

In this work we propose to learn a mapping from image pixels into a dense template grid through a fully convolutional network. We formulate this task as a regression problem and train our network by leveraging upon manually annotated facial landmarks “in-the-wild". We use such landmarks to establish a dense correspondence field between a three-dimensional object template and the input image, which then serves as the ground-truth for training our regression system. We show that we can combine ideas from semantic segmentation with regression networks, yielding a highly-accurate ‘quantized regression’ architecture.

Our system, called DenseReg allows us to estimate dense image-to-template correspondences in a fully convolutional manner. As such our network can provide useful correspondence information as a stand-alone system, while when used as an initialization for Statistical Deformable Models we obtain landmark localization results that largely outperform the current state-of-the-art on the challenging 300W benchmark. We thoroughly evaluate our method on a host of facial analysis tasks, and also demonstrate its use for other correspondence estimation tasks, such as modelling of the human ear. This work was publised in [22].

7.27. Structured Output Prediction and Learning for Deep Monocular 3D Human Pose Estimation

Participants: Stefan Kinauer, Riza Alp Guler, Siddhartha Chandra, Iasonas Kokkinos

In this work we address the problem of estimating 3D human pose froma single RGB image by blending a feed-forward Convolutional Neural Network (CNN) with a graphical model thatcouples the 3D positions of parts. The CNN populates a volumetric output space that represents the possible positions of 3D human joints, and also regresses the estimated displacements between pairs of parts. These constitute the ‘unary’ and ‘pairwise’ terms of the energy of a graphical model that resides in a 3D label spaceand delivers an optimal 3D pose configuration at its output. The CNN is trained onthe 3D human pose dataset 3.6M, the graphical model is trained jointly with the CNN in an end-to-end manner, allowing us to exploit both the discriminative powerof CNNs and the top-down information pertaining to human pose. We introduce(a) memory efficient methods for getting accurate voxel estimates for parts byblending quantization with regression (b) employ efficient structured prediction algorithms for 3D pose estimation using branch-and-bound and (c) develop a framework for qualitative and quantitative comparison of competing graphical models. We evaluate our work on the Human 3.6M dataset, demonstrating that exploiting the structure of the human pose in 3D yields systematic gains.

7.28. Newton-type Methods for Inference in Higher-Order Markov Random Fields

Participants: Hariprasad Kannan, Nikos Paragios

Linear programming relaxations are central to MAP inference in discrete Markov Random Fields. The ability to properly solve the Lagrangian dual is a critical component of such methods. In this paper, we study the benefit of using Newton-type methods to solve the Lagrangian dual of a smooth version of the problem. We investigate their ability to achieve superior convergence behavior and to better handle the ill-conditioned nature of the formulation, as compared to first order methods. We show that it is indeed possible to efficiently apply a trust region Newton method for a broad range of MAP inference problems. In this paper we propose a provably convergent and efficient framework that includes (i) excellent compromise between computational complexity and precision concerning the Hessian matrix construction, (ii) a damping strategy that aids efficient optimization , (iii) a truncation strategy coupled with a generic pre-conditioner for Conjugate Gradients, (iv) efficient sum-product computation for sparse clique potentials. Results for higher-order Markov Random Fields demonstrate the potential of this approach. This work was published in [23].

7.29. Alternating Direction Graph Matching

Participants: D. Khuê Lê-Huu, Nikos Paragios

In this work, we introduce a graph matching method that can account for constraints of arbitrary order, with arbitrary potential functions. Unlike previous decomposition approaches that rely on the graph structures, we introduce a decomposition of the matching constraints. Graph matching is then reformulated as a non-convex non-separable optimization problem that can be split into smaller and much-easier-to-solve subproblems, by means of the alternating direction method of multipliers. The proposed framework is modular, scalable, and can be instantiated into different variants. Two instantiations are studied exploring pairwise and higher-order constraints. Experimental results on widely adopted benchmarks involving synthetic and real examples demonstrate that the proposed solutions outperform existing pairwise graph matching methods, and competitive with the state of the art in higher-order settings. This work was published in [24].

7.30. Prediction and classification in biological and information networks

Participants: Fragkiskos Malliaros (in collaboration with Duong Nguyen, UC San Diego)
We investigate how network representation learning algorithms can be applied to deal with the problem of link prediction and classification in protein-protein interaction networks as well as in social and information networks. In particular, we have proposed BiasedWalk, a scalable, unsupervised feature learning algorithm that is based on biased random walks to sample context information about each node in the network.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Contract with General Electric Healthcare
Project title: Optimization methods for breast tomosynthesis
Duration: 2017-2020
Responsible: J.-C. Pesquet

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

Program: ANR Blanc International

- Project acronym: ADAMANTIUS
- Project title: Automatic Detection And characterization of residual Masses in pAtients with lymphomas through fusioN of whole-body diffusion-weighTed mrI on 3T and 18F-flUorodeoxyglucoSe pet/ct
- Duration: 9/2012-8/2015
- Coordinator: CHU Henri Mondor - FR

Program: ANR JCJC

- Project acronym: HICORE
- Project title: HIerarchical COmpositional REpresentations for Computer Vision
- Duration: 10/2010-9/2014
- Coordinator: ECP - FR

Program: ANR JCJC

- Project acronym: LearnCost
- Project title: Learning Model Constraints for Structured Prediction

Duration: 2014-2018

Coordinator: Inria Saclay - FR

Program: ANR JCJC

Project acronym: MajIC

Project title: Majorization-Minimization Algorithms for Image Computing

Duration: 2017-2021

Coordinator: E. Chouzenoux

Program: ITMOs Cancer & Technologies pour la santé d'Aviesan / INCa

Project acronym: CURATOR

Project title: Slice-to-Image Deformable Registration towards Image-based Surgery Navigation & Guidance

Duration: 12/2013-11/2015

Coordinator: ECP - FR

9.1.2. Others

Program: CNRS MASTODONS
Projet acronym: TABASCO
Project title: Traitement du bruit non Gaussien en spectroscopie
Duration: 2016-2018
Coordinator: E. Chouzenoux

Program: CNRS-CEFIPRA
Project acronym: NextGenBP
Project title: Looking Beyond Backpropagation in Deep Learning
Duration : 2017-2019
Coordinator: E. Chouzenoux

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

9.2.1.1. MOBOT

Title: Intelligent Active MObility Aid RoBOT integrating Multimodal Communication

Programm: FP7

Duration: February 2013 - January 2016

Coordinator: Technische Universität München

Partners:

Bartlomiej Marcin Stanczyk (Poland)

Athena Research and Innovation Center in Information Communication & Knowledge Technologies (Greece)

Bethanien Krankenhaus - Geriatrisches Zentrum - Gemeinnutzige (Germany)

Diaplasis Rehabilitation Center (Greece)

Ecole Centrale des Arts et Manufactures (France)

Institute of Communication and Computer Systems (Greece)

Technische Universitaet Muenchen (Germany)

Ruprecht-Karls-Universitaet Heidelberg (Germany)

Inria contact: Iasonas Kokkinos

Mobility disabilities are prevalent in our ageing society and impede activities important for the independent living of elderly people and their quality of life. The MOBOT project aims at supporting mobility and thus enforcing fitness and vitality by developing intelligent active mobility assistance robots for indoor environments that provide user-centred, context-adaptive and natural support. Our driving concept envisions cognitive robotic assistants that act (a) proactively by realizing an autonomous and context-specific monitoring of human activities and by subsequently reasoning on meaningful user behavioural patterns, as well as (b) adaptively and interactively, by analysing multi-sensory and physiological signals related to gait and postural stability, and by performing adaptive compliance control for optimal physical support and active fall prevention. Towards these targets, a multimodal action recognition system will be developed to monitor, analyse and predict user actions with a high level of accuracy and detail. The main thrust of our approach will be the enhancement of computer vision techniques with modalities such as range sensor images, haptic information as well as command-level speech and gesture recognition. Data-driven multimodal human behaviour analysis will be conducted and behavioural patterns will be extracted. Findings will be imported into a multimodal human-robot communication system, involving both verbal and nonverbal communication and will be conceptually and systemically synthesised into mobility assistance models taking into consideration safety critical requirements. All these modules will be incorporated in a behaviour-based and context-aware robot control framework. Direct involvement of end-user groups will ensure that actual user needs are addressed. Finally, user trials will be conducted to evaluate and benchmark the overall system and to demonstrate the vital role of MOBOT technologies for Europe's service robotics.

9.2.1.2. Strategie

Title: Statistically Efficient Structured Prediction for Computer Vision and Medical Imaging

Programm: FP7

Duration: January 2014 - December 2017

Coordinator: Inria

Inria contact: Matthew Blaschko

'Inference in medical imaging is an important step for disease diagnosis, tissue segmentation, alignment with an anatomical atlas, and a wide range of other applications. However, imperfections in imaging sensors, physical limitations of imaging technologies, and variation in the human population mean that statistical methods are essential for high performance. Statistical learning makes use of human provided ground truth to enable computers to automatically make predictions on future examples without human intervention. At the heart of statistical learning methods is risk minimization - the minimization of the expected loss on a previously unseen image. Textbook methods in statistical learning are not generally designed to minimize the expected loss for loss functions appropriate to medical imaging, which may be asymmetric and non-modular. Furthermore, these methods often do not have the capacity to model interdependencies in the prediction space, such as those arising from spatial priors, and constraints arising from the volumetric layout of human anatomy. We aim to develop new statistical learning methods that have these capabilities, to develop efficient learning algorithms, to apply them to a key task in medical imaging (tumor segmentation), and to prove their convergence to optimal predictors. To achieve this, we will leverage the structured prediction framework, which has shown impressive empirical results on a wide range of learning tasks. While theoretical results giving learning rates are available for some algorithms, necessary and sufficient conditions for consistency are not known for structured prediction. We will consequently address this issue, which is of key importance for algorithms that will be applied to life critical applications, e.g. segmentation of brain tumors that will subsequently be targeted by radiation therapy or removed by surgery. Project components will address both theoretical and practical issues.'

9.2.2. I-SUPPORT

Title: ICT-Supported Bath Robots

Project-Team GALEN 17

Program: FP7

Duration: March 2015 - March 2018

Coordinator: Robotnik Automation S.L.L.

Partners:
Bethanien Krankenhaus - Geriatrisches Zentrum - Gemeinnutzige GMBH (Germany)
Fondazione Santa Lucia (Italy)
Institute of Communication and Computer Systems (Greece)
Karlsruher Institut für Technologie (Germany)
Theofanis Alexandridis Kai Sia Ee (OMEGATECH) (Greece)
Robotnik Automation Sll (Spain)
Scuola Superiore di Studi Universitari E di Perfezionamento Sant'Anna (Italy)
Frankfurt University of Applied Sciences (Germany)

Inria contact: Iasonas Kokkinos

The I-SUPPORT project envisions the development and integration of an innovative, modular, ICTsupported service robotics system that supports and enhances older adults' motion and force abilities and assists them in successfully, safely and independently completing the entire sequence of bathing tasks, such as properly washing their back, their upper parts, their lower limbs, their buttocks and groin, and to effectively use the towel for drying purposes. Advanced modules of cognition, sensing, context awareness and actuation will be developed and seamlessly integrated into the service robotics system to enable the robotic bathing system to adapt to the frail elderly population' capabilities and the frail elderly to interact in a master-slave mode, thus, performing bathing activities in an intuitive and safe way. Adaptation and integration of state-of-the-art, cost-effective, soft-robotic manipulators will provide the hardware constituents, which, together with advanced human-robot force/compliance control that will be developed within the proposed project, will form the basis for a safe physical human-robot interaction that complies with the most up-to-date safety standards. Human behavioural, sociological, safety, ethical and acceptability aspects, as well as financial factors related to the proposed service robotic infrastructure will be thoroughly investigated and evaluated so that the I-SUPPORT end result is a close-to-market prototype, applicable to realistic living settings.

9.3. International Initiatives

9.3.1. Inria International Partners

9.3.1.1. Informal International Partners

Sup'Com Tunis - Collaborative research with Amel Benazza-Benhayia. Collaboration Topic: Multi-spectral imaging.

Universidad Tecnica Federico Santa Maria - Collaborative research with Luis M. Briceno Arias. Collaboration Topics: Variational approaches for monotone inclusions.

University of Patras, Greece - Collaborative research with V. Megalooikonomou. Collaboration Topic: Biosignal analysis.

University of Pennsylvania - Collaborative research with Aristeidis Sotiras. Collaboration Topic: Higher Order Graphs in biomedical image analysis.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

9.4.1.1. Internships

The following international students did an internship at CVN in the past year:

Huidong Liu, Stony Brook University, NY (may 2017)
Zhixin Shu, Stony Brook University, NY (may 2017)
Vu Nguyen, Stony Brook University, NY (jul. 2017)
Han Anh Vu Le, Houston University (jul. 2017)
Anisia Florescu, University of Galati Romania (feb. 2017)
Vyacheslav Dudar, Taras Sheuchenko National University of Kyiv (nov. 2017)
Carla Bertolocchi, Universita degli studi di Modena e Reggio Emilia (dec. 2017)
Yana Vedel, Taras Sheuchenko National University of Kyiv (dec. 2017)

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. Member of the Organizing Committees

E. Chouzenoux (co-organizer). Organization of the special session "Interactions entre méthodes d'optimisation et algorithmes de simulation stochastique" in the conference GRETSI 2017, Juan-les-Pins, 5-8 sep. 2017.

E. Chouzenoux and J.C. Pesquet. Organization of the special session "Advanced optimization methods for solving inverse problems at a large scale" in the International Biomedical and Astronomical Signal Processing Frontiers workshop (BASP 2017), Villars-sur-Ollon, Suisse, 29 jan.- 3 feb. 2017.

E. Belilovsky. Co-Organizer of Learning with Limited Labeled Data Workshop, in conjunction with Conference on Neural Information Processing Systems (NIPS).

10.1.2. Scientific Events Selection

10.1.2.1. Member of the Conference Program Committees

E. Chouzenoux. Member of the technical committee "Signal Processing Theory and Methods" of the IEEE Signal Processing Society.

F. Malliaros. Member of the program committee at: International Conference on World Wide Web (WWW), Conference on Artificial Intelligence (AAAI), International Joint Conference on Artificial Intelligence (IJCAI).

10.1.2.2. Reviewer

The members of the team reviewed numerous papers for several international conferences, such as for the annual conferences on Computer Vision and Pattern Recognition (CVPR), Medical Image Computing and Computer Assisted Intervention (MICCAI), Neural Information Processing Systems (NIPS) and International Conference on Learning Representations (ICLR), IEEE International Conference and Acoustics Speech and Signal Processing (ICASSP), IEEE International Conference on Image Processing (ICIP), IEEE Statistical Signal Processing workshop (SSP), European Signal Processing Conference (EUSIPCO).

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

Zacharaki, Evangelia I.: Medical Physics (guest editor), International Journal of Radiology, Dataset Papers in Science (Radiology), International Journal of Biostatistics & Computational Biology

10.1.3.2. Reviewer - Reviewing Activities

E. I. Zacharaki: IEEE Trans. on Medical Imaging (T-MI), Medical Image Analysis (MedIA), Trans. on Biomedical Engineering, Neuroimage, Artificial Intelligence in Medicine, Expert Systems with Applications, Bioinformatics, European Radiology

M. Vakalopoulou: International Journal of Computer Assisted Radiology and Surgery (IJCARS), IEEE Trans. on Geoscience and Remote Sensing (TGRS), Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS), ISPRS Journal of Photogrammetry and Remote Sensing, Computer Methods and Programs in Biomedicine (CMPB), Pattern Recognition Letters

E. Chouzenoux: IEEE Trans. on Image Processing, IEEE Trans. Signal Processing, SIAM Journal on Imaging Science, Journal of Optimization Theory and Applications, Journal of Global Optimization.

J.-C. Pesquet: IEEE Trans. on Signal Processing, IEEE Trans. on Image Processing, IEEE Trans. on Information Theory (IEEE-TI), Signal Processing, SIAM Journal on Optimization, SIAM Journal on Imaging Sciences, Journal of Mathematical Imaging and Vision, Journal of Optimization Theory and Applications.

F. Malliaros: Data Mining and Knowledge Discovery (DAMI), ACM Transactions on Knowledge Discovery from Data (TKDD).

10.1.4. Invited Talks

J.-C. Pesquet:
Villars sur Oulon (International Biomedical and Astronomical Signal Processing Frontiers workshop, Feb. 2017)
Taras Shevchenko National University of Kiev (Feb. 2017)
Politecnico di Milano (Apr. 2017)
Gdansk University (Mar. 2017)
Juans-les-Pins (Colloque GRETSI, Sep. 2017)
Polytechnic University of Warsaw (Oct. 2017)
Université Libre de Bruxelles (Nov. 2017)

E. Chouzenoux:
Invited seminar, LCS, Saclay, 8 décembre 2017.
Invited seminar, IFPEN, Rueil Malmaison, 6 décembre 2017.
Invited seminar, S3, CentraleSupélec, 24 novembre 2017.
Invited seminar, IIIT Delhi, 8 novembre 2017.
Journées annuelles 2017 des GdR MOA et MIA, Talence, 18 octobre 2017.
Journées Franco-Chiliennes d'Optimisation, Toulouse, 6 juillet 2017
Ecole d'été "Structured Regularization for High-Dimensional Data Analysis", IHP, 21 juin 2017.
Invited seminar, Université de Modène, 3 avril 2017.
Journée GDR ISIS, Inversion et problèmes multi-*, IAP, Paris, 20 mars 2017.

10.1.5. Leadership within the Scientific Community

J.-C. Pesquet is senior member of the Institut Universitaire de France and a Fellow of IEEE.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master : Corbineau, Marie-Caroline and Pesquet, Jean-Christophe. Advanced course on Optimization, 33h, M1, CentraleSupélec, FR

Master : Chouzenoux Emilie and Pesquet, Jean-Christophe. Foundations of Distributed and Large Scale Computing, 26h, M.Sc. in Data Sciences and Business Analytics, CentraleSupélec and ESSEC Business School, FR

Master: Zacharaki, Evangelia I. Foundations in Machine Learning, 36, M2 DataScience, CentraleSupélec , FR

Master: Malliaros, Fragkiskos. Machine Learning, 30h, M.Sc. in Data Sciences and Business Analytics, CentraleSupélec and ESSEC Business School, FR

Master: Kannan, Hariprasad and Sahasrabudhe, Mihir: Programming and Languages, 24h, M.Sc. in Data Science and Business Analytics, CentraleSupélec and ESSEC Business School, FR

10.2.2. Supervision

PhD in progress : Mihir Sahasrabudhe, Understanding Correlations in High-Dimensional Spaces and their Applications in Medical Imaging and Computer Vision, 2015-2019, Nikos Paragios

PhD in progress : Siddhartha Chandra, Deep Structured Prediction for Dense Labeling Tasks in Computer Vision, 2014-2018, Iasonas Kokkinos

PhD in progress : Marie-Caroline Corbineau, Fast online optimization algorithms for machine learning and medical imaging, 2016-2019, supervised by Emilie Chouzenoux and J.-C. Pesquet

PhD in progress : Loubna El Gueddari, Parallel proximal algorithms for compressed sensing MRI reconstruction - Applications to ultra-high magnetic field imaging, 2016-2019, supervised by J.-C. Pesquet and Ph. Ciuciu (Inria PARIETAL)

PhD in progress: Arthur Marmin, Rational models optimized exactly for chemical processes improvement, supervised by Marc Castella (Telecom Paristech) and J.-C. Pesquet

PhD in progress: Maïssa Sghaier, clinical Task-Based Reconstruction in tomosynthesis, supervised by E. Chouzenoux, J.-C. Pesquet and G. Palma (GE Healthcare)

PhD in progress : Evgenios Kornaropoulos, Diffusion Coefficient: a novel computer aided biomarker, 2013-2017, N. Paragios

PhD in progress: Guillaume Chassagnon, Development of new quantitative imaging biomarkers for obstructive and interstitial lung diseases, 2016-2019, N. Paragios

PhD in progress: Roger Sun,Deep learning and computer vision approaches on medical imaging and genomic data to improve the prediction of anticancer therapies' efficacy, 2017-2020, N. Paragios

PhD in progress: Théo Estienne, Improving anticancer therapies efficacy through Machine Learning on Medical Imaging & Genomic Data, 2017-2020, N. Paragios

PhD in progress: Abdulkadir Celikkanat, Representation learning methods on graphs, 2017-2020, N. Paragios and F. Malliaros

10.2.3. Juries

The faculty members of the team (N. Paragios, J.-C. Pesquet, E. Chouzenoux and F. Malliaros) participated to numerous PhD Thesis Committees, HDR Committees and served as Grant Reviewers.
Emilie Chouzenoux, Mihir Sahasrabudhe, Jean-Christophe Pesquet, Marie-Caroline Corbineau and Nikos Paragios were part of the jury committee for several end-of-course internship presentations in CentraleSupélec.

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] E. CHOUZENOUX.*Algorithmes de majoration-minimisation. Application aux problèmes inverses de grande taille en signal/image*, Université Paris Est - Marne-la-Vallée, December 2017, Habilitation à diriger des recherches, https://hal.archives-ouvertes.fr/tel-01661236.

[2] E. N. KORNAROPOULOS.*Deformable Group-wise Image Registration for Motion Estimation in 4D Medical Imaging* , Ecole Centrale Paris, June 2017, https://hal.inria.fr/tel-01577683.

[3] E. I. ZACHARAKI.*Computational methods towards image-based biomarkers and beyond*, Université Paris-Est, March 2017, Habilitation à diriger des recherches, https://hal.inria.fr/tel-01648583.

Articles in International Peer-Reviewed Journal

[4] F. ABBOUD, E. CHOUZENOUX, J.-C. PESQUET, J.-H. CHENOT, L. LABORELLI.*Dual Block Coordinate Forward-Backward Algorithm with Application to Deconvolution and Deinterlacing of Video Sequences*, in "Journal of Mathematical Imaging and Vision", November 2017, vol. 59, n^{o} 3, p. 415-431 [*DOI :* 10.1007/S10851-016-0696-Y], https://hal.archives-ouvertes.fr/hal-01418393.

[5] S. ALCHATZIDIS, A. SOTIRAS, E. I. ZACHARAKI, N. PARAGIOS.*A Discrete MRF Framework for Integrated Multi-Atlas Registration and Segmentation*, in "International Journal of Computer Vision", January 2017 [*DOI :* 10.1007/S11263-016-0925-2], https://hal.archives-ouvertes.fr/hal-01359094.

[6] S. AMIDI, A. AMIDI, D. VLACHAKIS, N. PARAGIOS, E. I. ZACHARAKI.*Automatic single- and multi-label enzymatic function prediction by machine learning*, in "PeerJ", 2017, vol. 5, p. 1-16 [*DOI :* 10.7717/PEERJ.3095], https://hal.inria.fr/hal-01648529.

[7] A. CHERNI, E. CHOUZENOUX, M.-A. DELSUC.*PALMA, an improved algorithm for DOSY signal processing*, in "Analyst", 2017, vol. 142, n^{o} 5, p. 772 - 779 [*DOI :* 10.1039/C6AN01902A], https://hal.archives-ouvertes.fr/hal-01613209.

[8] E. CHOUZENOUX, J.-C. PESQUET.*A Stochastic Majorize-Minimize Subspace Algorithm for Online Penalized Least Squares Estimation*, in "IEEE Transactions on Signal Processing", September 2017, vol. 65, n^{o} 18, p. 4770 - 4783 [*DOI :* 10.1109/TSP.2017.2709265], https://hal.archives-ouvertes.fr/hal-01613204.

[9] E. FERRANTE, N. PARAGIOS.*Graph-Based Slice-to-Volume Deformable Registration*, in "International Journal of Computer Vision", 2017 [*DOI :* 10.1007/S11263-017-1040-8], https://hal.inria.fr/hal-01576314.

[10] E. FERRANTE, N. PARAGIOS.*Slice-to-volume medical image registration: A survey*, in "Medical Image Analysis", July 2017, vol. 39, p. 101 - 123 [*DOI :* 10.1016/J.MEDIA.2017.04.010], https://hal.inria.fr/hal-01650929.

[11] V. G. KANAS, E. I. ZACHARAKI, G. A. THOMAS, P. O. ZINN, V. MEGALOOIKONOMOU, R. R. COLEN.*Learning MRI-based classification models for MGMT methylation status prediction in glioblastoma*,

in "Computer Methods and Programs in Biomedicine", January 2017 [*DOI :* 10.1016/J.CMPB.2016.12.018], https://hal.inria.fr/hal-01423323.

[12] E. J. LIMKIN, R. SUN, L. DERCLE, E. I. ZACHARAKI, C. ROBERT, S. REUZÉ, A. SCHERNBERG, N. PARAGIOS, E. DEUTSCH, C. FERTÉ.*Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology*, in "Annals of Oncology", June 2017, vol. 28, n° 6, p. 1191 - 1206 [*DOI :* 10.1093/ANNONC/MDX034], https://hal.inria.fr/hal-01648559.

[13] Y. MARNISSI, Y. ZHENG, E. CHOUZENOUX, J.-C. PESQUET.*A Variational Bayesian Approach for Image Restoration. Application to Image Deblurring with Poisson-Gaussian Noise*, in "IEEE Transactions on Computational Imaging", 2017, 16 [*DOI :* 10.1109/TCI.2017.2700203], https://hal.archives-ouvertes.fr/hal-01613200.

[14] E. PIPPA, V. G. KANAS, E. I. ZACHARAKI, V. TSIRKA, M. KOUTROUMANIDIS, V. MEGALOOIKONOMOU.*EEG-based Classification of Epileptic and Non-epileptic Events using Multi-array Decomposition*, in "International Journal of Monitoring and Surveillance Technologies Research", January 2017, https://hal.archives-ouvertes.fr/hal-01359125.

[15] E. PIPPA, E. I. ZACHARAKI, M. KOUTROUMANIDIS, V. MEGALOOIKONOMOU.*Data fusion for paroxysmal events' classification from EEG*, in "Journal of Neuroscience Methods", January 2017, vol. 275, p. 55-65 [*DOI :* 10.1016/J.JNEUMETH.2016.10.004], https://hal.inria.fr/hal-01426373.

[16] A. PIRAYRE, C. COUPRIE, L. DUVAL, J.-C. PESQUET.*BRANE Clust: Cluster-Assisted Gene Regulatory Network Inference Refinement*, in "IEEE/ACM Transactions on Computational Biology and Bioinformatics", 2017, https://hal-ifp.archives-ouvertes.fr/hal-01330638.

[17] R. SUN, E. J. LIMKIN, L. DERCLE, S. REUZÉ, E. I. ZACHARAKI, C. CHARGARI, A. SCHERNBERG, A.-S. DIRAND, A. ALEXIS, N. PARAGIOS, E. DEUTSCH, C. FERTÉ, C. ROBERT.*Computational medical imaging (radiomics) and potential for immuno-oncology*, in "Cancer Radiothérapie", August 2017, vol. 21, n° 6-7, p. 648-654 [*DOI :* 10.1016/J.CANRAD.2017.07.035], https://hal.inria.fr/hal-01668902.

[18] E. I. ZACHARAKI.*Prediction of protein function using a deep convolutional neural network ensemble*, in "PeerJ Computer Science", 2017, vol. 3, p. 1-17 [*DOI :* 10.7717/PEERJ-CS.124], https://hal.inria.fr/hal-01648534.

International Conferences with Proceedings

[19] E. B. BELILOVSKY, K. KASTNER, G. VAROQUAUX, M. B. BLASCHKO.*Learning to Discover Sparse Graphical Models*, in "International Conference on Machine Learning", Sydney, Australia, August 2017, https://arxiv.org/abs/1605.06359 , https://hal.inria.fr/hal-01306491.

[20] G. CHIERCHIA, A. CHERNI, E. CHOUZENOUX, J.-C. PESQUET.*Approche de Douglas-Rachford aléatoire par blocs appliquée à la régression logistique parcimonieuse*, in "GRETSI 2017", Juan les Pins, France, Actes du 26e colloque GRETSI, September 2017, p. 1-4, https://hal.archives-ouvertes.fr/hal-01634525.

[21] E. FERRANTE, P. K. DOKANIA, R. MARINI, N. PARAGIOS.*Deformable Registration through Learning of Context-Specific Metric Aggregation*, in "Machine Learning in Medical Imaging Worlshop. MLMI (MICCAI 2017)", Quebec City, Canada, September 2017, https://arxiv.org/abs/1707.06263 - Accepted for publication in

the 8th International Workshop on Machine Learning in Medical Imaging (MLMI 2017), in conjunction with MICCAI 2017, https://hal.inria.fr/hal-01650956.

[22] R. A. GULER, G. TRIGEORGIS, E. ANTONAKOS, P. SNAPE, S. ZAFEIRIOU, I. KOKKINOS.*DenseReg: Fully Convolutional Dense Shape Regression In-the-Wild*, in "2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)", Honolulu, United States, IEEE, CVF, July 2017, p. 6799-6808, https://arxiv.org/abs/1612.01202 , https://hal.archives-ouvertes.fr/hal-01637896.

[23] H. KANNAN, N. KOMODAKIS, N. PARAGIOS.*Newton-type Methods for Inference in Higher-Order Markov Random Fields*, in "IEEE International Conference on Computer Vision and Pattern Recognition", Honolulu, United States, July 2017, p. 7224 - 7233, https://hal.archives-ouvertes.fr/hal-01580862.

[24] D. K. LÊ-HUU, N. PARAGIOS.*Alternating Direction Graph Matching*, in "IEEE Conference on Computer Vision and Pattern Recognition (CVPR)", Honolulu, United States, July 2017, https://arxiv.org/abs/1611.07583 , https://hal.inria.fr/hal-01580824.

[25] E. OYALLON, E. BELILOVSKY, S. ZAGORUYKO.*Scaling the Scattering Transform: Deep Hybrid Networks*, in "International Conference on Computer Vision (ICCV)", Venice, Italy, October 2017, https://arxiv.org/abs/1703.08961 , https://hal.inria.fr/hal-01495734.

Conferences without Proceedings

[26] A. BENFENATI, E. CHOUZENOUX, J.-C. PESQUET.*A Proximal Approach for Solving Matrix Optimization Problems Involving a Bregman Divergence*, in "BASP 2017 - International Biomedical and Astronomical Signal Processing Frontiers workshop", villars-sur-oulon, Switzerland, January 2017, https://hal.archives-ouvertes.fr/hal-01613292.

[27] S. CADONI, E. CHOUZENOUX, J.-C. PESQUET, C. CHAUX.*A Block Parallel Majorize-Minimize Memory Gradient Algorithm*, in "BASP 2017 - International Biomedical and Astronomical Signal Processing Frontiers workshop", Villars-sur-Oulon, Switzerland, January 2017, 1, https://hal.archives-ouvertes.fr/hal-01634531.

[28] V. DUDAR, G. CHIERCHIA, E. CHOUZENOUX, J.-C. PESQUET, V. V. SEMENOV.*A Two-Stage Subspace Trust Region Approach for Deep Neural Network Training*, in "25th European Signal Processing Conference (EUSIPCO 2017)", Kos Island, Greece, August 2017, https://hal.archives-ouvertes.fr/hal-01634538.

[29] Q. WEI, E. CHOUZENOUX, J.-Y. TOURNERET, J.-C. PESQUET.*A Fast Algorithm Based on a Sylvester-like Equation for LS Regression with GMRF Prior*, in "CAMSAP 2017- IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing", Curaçao, Netherlands Antilles, December 2017, https://hal.archives-ouvertes.fr/hal-01635601.

Scientific Popularization

[30] S. CHANDRA, N. USUNIER, I. KOKKINOS.*Dense and Low-Rank Gaussian CRFs Using Deep Embeddings*, in "ICCV 2017 - International Conference on Computer Vision", Venice, Italy, September 2017, https://hal.inria.fr/hal-01646293.

Other Publications

[31] F. ABBOUD, E. CHOUZENOUX, J.-C. PESQUET, J.-H. CHENOT, L. LABORELLI.*An Alternating Proximal Approach for Blind Video Deconvolution*, December 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01668437.

[32] A. BENFENATI †, E. CHOUZENOUX, J.-C. PESQUET.*A proximal approach for a class of matrix optimization problems*, December 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01673027.

[33] L. BRICEÑO-ARIAS, G. CHIERCHIA, E. CHOUZENOUX, J.-C. PESQUET.*A Random Block-Coordinate Douglas-Rachford Splitting Method with Low Computational Complexity for Binary Logistic Regression*, December 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01672507.

[34] Y. MARNISSI, E. CHOUZENOUX, A. BENAZZA-BENYAHIA, J.-C. PESQUET.*An Auxiliary Variable Method for MCMC Algorithms in High Dimension*, December 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01661234.

[35] J. YU, M. BLASCHKO.*The Lovász Hinge: A Novel Convex Surrogate for Submodular Losses*, May 2017, working paper or preprint, https://hal.inria.fr/hal-01241626.

[36] Y. ZHENG, A. PIRAYRE, L. DUVAL, J.-C. PESQUET.*Joint restoration/segmentation of multicomponent images with variationalBayes and higher-order graphical models (HOGMep)*, May 2017, working paper or preprint, https://hal-ifp.archives-ouvertes.fr/hal-01528856.

Project-Team GAMMA3

Automatic mesh generation and advanced methods

IN PARTNERSHIP WITH:
Université de Technologie de Troyes

RESEARCH CENTER
Saclay - Île-de-France

THEME
Numerical schemes and simulations

Table of contents

Project-Team GAMMA3

Creation of the Project-Team: 2010 January 01

Keywords:

Computer Science and Digital Science:

A2.5. - Software engineering
A5.2. - Data visualization
A6.1. - Mathematical Modeling
A6.2. - Scientific Computing, Numerical Analysis & Optimization
A7.1. - Algorithms
A8.3. - Geometry, Topology

Other Research Topics and Application Domains:

B5.2.3. - Aviation
B5.2.4. - Aerospace

1. Personnel

Research Scientists

Paul-Louis George [Team leader, Inria, Senior Researcher]
Frederic Alauzet [Inria, Senior Researcher, HDR]
Patrick Laug [Inria, Senior Researcher, HDR]
Adrien Loseille [Inria, Researcher]
David Marcum [Inria, International Chair, Advanced Research Position]

Faculty Member

Houman Borouchaki [Univ de technologie de Troyes, Professor]

External Collaborators

Eléonore Gauci [Univ Pierre et Marie Curie, from Sep 2017]
David Marcum [Mississippi State University]
Loic Marechal [Inria]

PhD Students

Bastien Andrieu [ONERA]
Rémi Feuillet [École Nationale Supérieure de Techniques Avancées]
Loïc Frazza [Ecole polytechnique]
Julien Vanharen [Inria, May 2017]

Post-Doctoral Fellows

Olivier Coulaud [Inria, until Aug 2017]
Julien Vanharen [Inria, from Jun 2017]

Administrative Assistants

Jessica Gameiro [Inria]
Emmanuelle Perrot [Inria]

2. Overall Objectives

2.1. Introduction

Une branche importante des sciences de l'ingénieur s'intéresse aux calculs des solutions d'équations aux dérivées partielles très variées (en mécanique du solide, en mécanique des fluides, en modélisation de problèmes thermiques, ...) par la méthode des éléments ou des volumes finis. Ces méthodes utilisent comme support spatial des calculs un maillage du domaine sur lequel les équations sont formulées. Par suite, les algorithmes (de construction) de maillages occupent un rôle important dans toute simulation par la méthode des éléments ou des volumes finis d'un problème modélisé en équations aux dérivées partielles. En particulier, la précision, voire la validité, des solutions calculées est liée aux propriétés du maillage utilisé [12].

L'équipe-projet GAMMA3 a été créé en 2010 à la suite du projet GAMMA. L'équipe est bilocalisée avec une partie à l'UTT (Troyes) et l'autre à Rocquencourt puis Saclay. Les thèmes du projet regroupent un ensemble d'activités concernant les points indiqués ci-dessus, en particulier, l'aspect génération automatique de maillages afin de construire les supports utilisés par les méthodes d'éléments ou de volumes finis. Sont également étudiés les aspects de modélisation géométrique, de post-traitement et de visualisation des résultats issus de tels calculs [13].

L'évolution de la demande en termes de génération automatique de maillages implique une évolution des méthodes classiques de création de maillages vers des méthodes permettant de construire des maillages contrôlés. Les maillages doivent donc être soit isotropes, le contrôle portant sur des tailles souhaitées, soit anisotropes, le contrôle portant à la fois sur des directions et des tailles selon ces dernières.

Le développement d'algorithmes de maillages gouvernés sert de support naturel à la conception de boucles de maillages adaptatifs qui, via un estimateur d'erreurs *a posteriori*, permettent de contrôler la qualité des solutions. Les estimateurs d'erreurs sont issus d'applications en mécanique des fluides (Inria) et du solide (UTT). Leurs validations reposent sur le développement de solveurs avancés, en particulier, dans ces disciplines. Ces deux points (estimateurs et solveurs) constituent au moins la moitié de nos recherches.

Ces préoccupations amènent à considérer le problème du maillage des domaines de calculs en eux-mêmes tout comme celui du maillage ou du remaillage des courbes et surfaces, frontières de ces domaines.

La taille, en termes de nombre de nœuds, des maillages nécessaires pour certaines simulations, amène à travailler sur la parallélisation des processus de calculs. Cette problématique conduit également à s'intéresser à l'aspect multi-cœurs au niveau des algorithmes de maillages proprement dits.

Simultanément, le volume des résultats obtenus dans de telles simulations, nécessite d'envisager le post-traitement de ces résultats en parallèle ou par des méthodes appropriées.

Par ailleurs, de nombreux problèmes partent de saisies *scanner* (ou autre système discret) des géométries à traiter et demandent d'en déduire des maillages de surfaces aptes à être, par la suite, traités par les méthodes classiques (de remaillage, d'optimisation, de calculs).

Enfin, la maturité de certaines méthodes (victimes de leur succès) conduit les utilisateurs à demander plus et à considérer des problèmes de maillage ou des conditions d'utilisations extrêmes induisant des algorithmes *a priori* inattendus.

Les objectifs du projet GAMMA3 consistent à étudier l'ensemble des points mentionnés ci-dessus afin de rendre automatique le calcul de la solution d'un problème donné avec une précision imposée au départ. Par ailleurs, certaines des techniques utilisées dans les problématiques de maillage sont utilisables dans d'autres disciplines (compression d'images pour ne citer qu'un seul exemple).

3. New Software and Platforms

3.1. ABL4FLO

KEYWORDS: Boundary layers - Hybrid meshes

FUNCTIONAL DESCRIPTION: ABL4FLO is designed to generate 3D adapted boundary layer meshes by using a cavity-based operator.

- Participant: Adrien Loseille
- Contact: Adrien Loseille

3.2. AMA4FLO

Anisotropic Mesh Adaptation 4 FLOw
KEYWORDS: 3D - Mesh adaptation
FUNCTIONAL DESCRIPTION: 3D, surface, 2D anisotropic mesh generation

- Participant: Adrien Loseille
- Contact: Adrien Loseille
- URL: http://pages.saclay.inria.fr/adrien.loseille/index.php?page=softwares

3.3. BL2D

KEYWORDS: Abstraction - Meshing - Isotropic - Anisotropic - Delaunay
FUNCTIONAL DESCRIPTION: This software package stems from a former one called BL2D-V1. The meshing method is of controled Delaunay type, isotropic or anisotropic. The internal point generation follows a frontal logic, and their connection is realised as in a classical Delaunay approach. Quadrilaterals are obtained by a pairing process. The direct construction of degree 2 element has been made possible via the control of the domain boundary mesh, in order to ensure the desired compatibility. The boundary middle nodes are located according to the curvilinear abscissa. The internal middle nodes are, by default, at the middle of the corresponding edges.

RELEASE FUNCTIONAL DESCRIPTION: Par rapport à la version V1, il offre de nombreuses possibilités nouvelles : méthode frontale, triangles quadratiques courbes, quadrilatères de degré 1ou 2, frontières déformables, allocation dynamique de mémoire, etc

- Participants: Houman Borouchaki and Patrick Laug
- Contact: Patrick Laug
- URL: http://pages.saclay.inria.fr/patrick.laug/logiciels/bl2d-v2/INDEX.html

3.4. BL2D-ABAQ

KEYWORDS: Anisotropic - Delaunay - Automatic mesher
FUNCTIONAL DESCRIPTION: The meshing method is the same as BL2D in an adaptive process. An a posteriori error estimation of a solution at the nodes of the current mesh results in a size map. A new mesh staisfying these size specifications (made continuous is built, and the solution is interpolated on the new mesh.

- Participants: Abel Cherouat, Houman Borouchaki and Patrick Laug
- Contact: Patrick Laug

3.5. BLGEOL

KEYWORDS: Automatic mesher - Geologic structure

FUNCTIONAL DESCRIPTION: BLGEOL-V1 software can generate hex-dominant meshes of geologic structures complying with different geometric constraints: surface topography (valleys, reliefs, rivers), geologic layers and underground workings. First, a reference 2D domain is obtained by projecting all the line constraints into a horizontal plane. Different size specifications are given for rivers, outcrop lines and workings. Using an adaptive methodology, the size variation is bounded by a specified threshold in order to obtain a high quality quad-dominant mesh. Secondly, a hex-dominant mesh of the geological medium is generated by a vertical extrusion, taking into account the surfaces found (interfaces between two layers, top or bottom faces of underground workings). The generation of volume elements follows a global order established on the whole set of surfaces to ensure the conformity of the resulting mesh.

- Participants: Houman Borouchaki and Patrick Laug
- Contact: Patrick Laug
- URL: https://team.inria.fr/gamma3/project-presentation/gamma-software/

3.6. BLMOL

KEYWORDS: Mesher - Molecular surface
SCIENTIFIC DESCRIPTION: An increasingly important part of quantum chemistry is devoted to molecular surfaces. To model such a surface, each constituting atom is idealized by a simple sphere. Surface mesh generation techniques are then used either for visualization or for simulation, where mesh quality has a strong influence on solution accuracy. First, a boundary representation (B-rep) of the surface is obtained, i.e. a set of patches and the topological relations between them. Second, an appropriate parameterization and a metric map are computed for each patch. Third, meshes of the parametric domains are generated with respect to an induced metric map, using a combined advancing-front generalized-Delaunay approach. Finally these meshes are mapped onto the entire surface. Several application examples illustrate various capabilities of our method.
FUNCTIONAL DESCRIPTION: BLMOL is a molecular surface mesher.

- Participants: Houman Borouchaki and Patrick Laug
- Contact: Patrick Laug
- URL: http://pages.saclay.inria.fr/patrick.laug/logiciels/blmol/INDEX.html

3.7. BLSURF

KEYWORDS: Automatic mesher - Molecular surface
FUNCTIONAL DESCRIPTION: An indirect method for meshing parametric surfaces conforming to a user-specifiable size map is used. First, from this size specification, a Riemannian metric is defined so that the desired mesh is one with unit length edges with respect to the related Riemannian space (the so-called

- Participants: Houman Borouchaki and Patrick Laug
- Partner: Université de Technologie de Troyes
- Contact: Patrick Laug
- URL: https://team.inria.fr/gamma3/project-presentation/gamma-software/

3.8. FEFLOA-REMESH

KEYWORDS: Scientific calculation - Anisotropic - Mesh adaptation
FUNCTIONAL DESCRIPTION: FEFLOA-REMESH is intended to generate adapted 2D, surface and volume meshes by using a unique cavity-based operator. The metric-aligned or metric-orthogonal approach is used to generate high quality surface and volume meshes independently of the anisotropy involved.

- Participants: Adrien Loseille and Frédéric Alauzet
- Contact: Adrien Loseille
- URL: http://pages.saclay.inria.fr/adrien.loseille/index.php?page=softwares

3.9. GAMANIC 3D

KEYWORDS: Tetrahedral mesh - Delaunay - Anisotropic size and direction control - Automatic mesher
FUNCTIONAL DESCRIPTION: GAMANIC3D is a volume mesher governed by a (anisotropic) size and directional specification metric field.

- Participants: Adrien Loseille, Éric Saltel, Frédéric Alauzet, Frederic Hecht, Houman Borouchaki and Paul Louis George
- Contact: Paul Louis Georges
- URL: http://www.meshgems.com/volume-meshing.html

3.10. GAMHIC 3D

KEYWORDS: Tetrahedral mesh - Delaunay - Isotropic - Automatic mesher
FUNCTIONAL DESCRIPTION: GAMHIC3D is a volume mesher governed by a (isotropic) size specification metric field.

- Participants: Adrien Loseille, Éric Saltel, Frédéric Alauzet, Frederic Hecht, Houman Borouchaki and Paul Louis George
- Contact: Paul Louis George
- URL: http://www.meshgems.com/volume-meshing.html

3.11. GHS3D

KEYWORDS: Tetrahedral mesh - Delaunay - Automatic mesher
FUNCTIONAL DESCRIPTION: GHS3D is an automatic volume mesher

- Participants: Adrien Loseille, Éric Saltel, Frédéric Alauzet, Frederic Hecht, Houman Borouchaki and Paul Louis George
- Contact: Paul Louis George
- URL: http://www.meshgems.com/volume-meshing.html

3.12. HEXOTIC

KEYWORDS: 3D - Mesh generation - Meshing - Unstructured meshes - Octree/Quadtree - Multi-threading - GPGPU - GPU
FUNCTIONAL DESCRIPTION: Input: a triangulated surface mesh and an optional size map to control the size of inner elements.

Output: a fully hexahedral mesh (no hybrid elements), valid (no negative jacobian) and conformal (no dangling nodes) whose surface matches the input geometry.

The software is a simple command line that requires no knowledge on meshing. Its arguments are an input mesh and some optional parameters to control elements sizing, curvature and subdomains as well as some features like boundary layers generation.

- Participant: Loïc Maréchal
- Partner: Distene
- Contact: Loïc Maréchal
- URL: https://team.inria.fr/gamma3/project-presentation/gamma-software/hexotic/

3.13. Nimbus 3D

KEYWORDS: Surface reconstruction - Point cloud

FUNCTIONAL DESCRIPTION: Nimbus3D is a surface reconstruction method piece of software

- Participants: Houman Borouchaki and Paul Louis George
- Contact: Paul Louis George
- URL: http://www.meshgems.com/volume-meshing.html

3.14. VIZIR

Maillages Clés en Main pour la Simulation Numérique
KEYWORD: Mesh
FUNCTIONAL DESCRIPTION: VIZIR is intended to visualize and modify interactively simplicial, hybrid and high order curved meshes.

- Participants: Adrien Loseille, Alexis Loyer and Julien Castelneau
- Contact: Adrien Loseille
- URL: http://pages.saclay.inria.fr/adrien.loseille/index.php?page=softwares

3.15. Wolf

KEYWORD: Scientific calculation
FUNCTIONAL DESCRIPTION: Numerical solver for the Euler and compressible Navier-Stokes equations with turbulence modelling. ALE formulation for moving domains. Modules of interpolation, mesh optimisation and moving meshes. Wolf is written in C++, and may be later released as an opensource library. FELiScE was registered in July 2014 at the Agence pour la Protection des Programmes under the Inter Deposit Digital Number IDDN.FR.001.340034.000.S.P.2014.000.10000.

- Participants: Adrien Loseille and Frédéric Alauzet
- Contact: Frédéric Alauzet
- URL: https://www.rocq.inria.fr/gamma/Frederic.Alauzet/code_eng.html

3.16. Wolf-Bloom

KEYWORD: Scientific calculation
FUNCTIONAL DESCRIPTION: Wolf-Bloom is a structured boundary layer mesh generator using a pushing approach. It start from an existing volume mesh and insert a structured boundary layer by pushing the volume mesh. The volume mesh deformation is solved with an elasticity analogy. Mesh-connectivity optimizations are performed to control volume mesh element quality.

- Participants: Adrien Loseille, David Marcum and Frédéric Alauzet
- Contact: Frédéric Alauzet
- URL: https://www.rocq.inria.fr/gamma/Frederic.Alauzet/code_eng.html

3.17. Wolf-Elast

KEYWORD: Scientific calculation
FUNCTIONAL DESCRIPTION: Wolf-Elast is a linear elasticity solver using the P1 Finite-Element method. The Young and Poisson coefficient can be parametrized. The linear system is solved using the Conjugate Gradient method with the LUSGS preconditioner.

- Participants: Adrien Loseille and Frédéric Alauzet
- Contact: Frédéric Alauzet
- URL: https://www.rocq.inria.fr/gamma/Frederic.Alauzet/code_eng.html

3.18. Wolf-Interpol

KEYWORD: Scientific calculation
FUNCTIONAL DESCRIPTION: Wolf-Interpol is a tool to transfer scalar, vector and tensor fields from one mesh to another one. Polynomial interpolation (from order 2 to 4) or conservative interpolation operators can be used. Wolf-Interpol also extract solutions along lines or surfaces.

- Participants: Adrien Loseille and Frédéric Alauzet
- Contact: Frédéric Alauzet
- URL: https://www.rocq.inria.fr/gamma/Frederic.Alauzet/code_eng.html

3.19. Wolf-MovMsh

KEYWORD: Scientific calculation
FUNCTIONAL DESCRIPTION: Wolf-MovMsh is a moving mesh algorithm coupled with mesh-connectivity optimization. Mesh deformation is computed by means of a linear elasticity solver or a RBF interpolation. Smoothing and swapping mesh optimization are performed to maintain good mesh quality. It handles rigid bodies or deformable bodies, and also rigid or deformable regions of the domain.

- Participants: Adrien Loseille and Frédéric Alauzet
- Contact: Paul Louis George
- URL: https://www.rocq.inria.fr/gamma/Frederic.Alauzet/code_eng.html

3.20. Wolf-Nsc

KEYWORD: Scientific calculation
FUNCTIONAL DESCRIPTION: Wolf-Nsc is numerical flow solver solving steady or unsteady turbulent compressible Euler and Navier-Stokes equations. The available turbulent models are the Spalart-Almaras and the Menter SST k-omega. A mixed finite volume - finite element numerical method is used for the discretization. Second order spatial accuracy is reached thanks to MUSCL type methods. Explicit or implicit time integration are available. It also resolved dual (adjoint) problem and compute error estimate for mesh adaptation.

- Participants: Adrien Loseille and Frédéric Alauzet
- Contact: Frédéric Alauzet
- URL: https://www.rocq.inria.fr/gamma/Frederic.Alauzet/code_eng.html

3.21. Wolf-Spyder

KEYWORD: Scientific calculation
FUNCTIONAL DESCRIPTION: Wolf-Spyder is a metric-based mesh quality optimizer using vertex smoothing and edge/face swapping.

- Participants: Adrien Loseille and Frédéric Alauzet
- Contact: Frédéric Alauzet
- URL: https://www.rocq.inria.fr/gamma/Frederic.Alauzet/code_eng.html

4. New Results

4.1. Element metric, element quality and interpolation error metric

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

The metric of a simplex of $\mathbb{R}^d$ is a metric tensor (symmetric positive definite matrix) in which the element is unity (regular with unit edge lengths). This notion is related to the problem of interpolation error of a given field over a mesh. Let K be a simplex and let us denote by v_{ij} the vector joining vertex i and vertex j of K. The metric of K can be written as:

$$\mathcal{M} = \frac{d+1}{2}\Big(\sum_{i<j} v_{ij}\,{}^t v_{ij}\Big)^{-1},$$

where $v_{ij}\,{}^t v_{ij}$ is a $d \times d$ rank 1 matrix related to edge ij.

The metric of a simplex also characterizes the element shape. In particular, if it is the identity, the element is unity. Hence, to define the shape quality of an element, one can determine the gap of the element metric $\mathcal{M}$ and the identity using different measures based on the eigenvalues $\lambda_i = \frac{1}{h_I^2}$ of $\mathcal{M}$ or those of $\mathcal{M}^{-1}$, e.g. h_i^2. Notice that metric $\mathcal{M}^{-1}$ is directly related to the geometry of the element (edge length, facet area, element volume). The first algebraic shape quality measure ranging from 0 to 1 is defined as the ratio of the geometric average of the eigenvalues of $\mathcal{M}^{-1}$ and their arithmetic average:

$$q(K) = \frac{\left(\prod_i h_i^2\right)^{\frac{1}{d}}}{\frac{1}{d}\sum_{i=1}^{d} h_i^2} = d\,\frac{\left(det(\mathcal{M}^{-1})\right)^{\frac{1}{d}}}{tr(\mathcal{M}^{-1})}.$$

As the geometric average is smaller than the arithmetic average, this measure is well defined. In addition, it is the algebraic reading of the well-known quality measure defined by:

$$q^{\frac{d}{2}}(K) = (d!)d^{\frac{d}{2}}(d+1)^{\frac{d-1}{2}}\frac{|K|}{\left(\sum_{i<j} l_{ij}^2\right)^{\frac{d}{2}}},$$

where the volume and the square of the edge lengths are involved. The algebraic meaning justifies the above geometric measure. The second algebraic shape quality measure is defined as the ratio of the harmonic average of the eigenvalues of $\mathcal{M}^{-1}$ and their arithmetic average (ranging also from 0 to 1):

$$q(K) = \frac{\left\{\frac{1}{d}\sum_{i=1}^{d}\frac{1}{h_i^2}\right\}^{-1}}{\frac{1}{d}\sum_{i=1}^{d} h_i^2} = \frac{d^2}{tr(\mathcal{M})tr(\mathcal{M}^{-1})}.$$

As above, this measure is well defined, the harmonic average being smaller the arithmetic one. From this measure, one can derive another well-known measure involving the roundness and the size of an element (measure which is widely used for convergene issues in finite element methods).

Note that these measures use the invariants of $\mathcal{M}^{-1}$ or $\mathcal{M}$ and thus can be evaluated from the coefficients of the characteristical polynomial of those matrices (avoiding the effective calculation of their eigenvalues). Another advantage of the above algebraic shape measures is their easy extensions in an arbitrary Euclidean space. Indeed, if $\mathcal{E}$ is the metric of such a space, the algebraic shape measures read:

$$q_{\mathcal{E}}(K) = d\,\frac{\left(det(\mathcal{M}^{-1}\mathcal{E})\right)^{\frac{1}{d}}}{tr(\mathcal{M}^{-1}\mathcal{E})} \quad , \quad q_{\mathcal{E}}(K) = \frac{d^2}{tr(\mathcal{E}^{-1}\mathcal{M})tr(\mathcal{M}^{-1}\mathcal{E})}.$$

Following this notion of a element metric, a natural work was done regarding how to define the element metric so as to achieve a given accuracy for the interpolation error of a function using a finite element approximation by means of simplices of arbitrary degree.

This is a new approach for the majoration of the interpolation error of a polynomial function of arbitrary degree n interpolated by a polynomial function of degree $n-1$. From that results a metric, the so-called interpolation metric, which allows for a control of the error. The method is based on the geometric and algebraic properties of the metric of a given element, metric in which the element is regular and unit. The interpolation metric plays an important role in advanced computations based on mesh adaptation. The method relies in a Bezier reading of the functions combined with Taylor expansions. In this way, the error in a given element is fully controled at the time the edges of the element are controled.

It is shown that the error in bounded as

$$|e(X)| \leq C\sum_{i<j} f^{(n)}(.)(v_{ij}, v_{ij}, ..., v_{ij}),$$

where C is a constant depending on d and n, v_{ij} is the edge from the vertices of K of index i and j, $f^{(n)}(.)$ is the derivative of order n of f applied to a n-uple uniquely composed of v_{ij}. If we consider the case $d = 2$ and $u = (x, y)$ is a vector in $\mathbb{R}^2$, we have

$$f^{(n)}(.)(u, u, ..., u) = \sum_{i=0}^{n-2} x^{n-2-i} y^i \; {}^t u \left(C_i^{n-2} \mathcal{H}_{(n-2,n-2-i,i)}\right) u \;,$$

where the quadratic forms $\mathcal{H}_{(n-2,n-2-i,i)}$ are defined by the matrices of order 2 (with constant entries):

$$\mathcal{H}_{(n-2,n-2-i,i)} = \begin{pmatrix} \frac{\partial^{(n)} f}{\partial x_1^{n-i} \partial x_2^{i}} & \frac{\partial^{(n)} f}{\partial x_1^{n-1-i} \partial x_2^{i+1}} \\ \\ \frac{\partial^{(n)} f}{\partial x_1^{n-1-i} \partial x_2^{i+1}} & \frac{\partial^{(n)} f}{\partial x_1^{n-2-i} \partial x_2^{i+2}} \end{pmatrix},$$

those matrices being the hessians of the derivatives of f of order $n-2$.

This work resulted in a paper submitted in a journal and currently under revision.

4.2. Realistic modeling of fractured geologic media

Participants: Patrick Laug [correspondant], Géraldine Pichot.

This study, in collaboration with project-team Serena, aims to model, in a realistic and efficient manner, natural fractured media. These media are characterized by their diversity of structures and organizations. Numerous studies in the past decades have evidenced the existence of characteristic structures at multiple scales. At fracture scale, the aperture distribution is widely correlated and heterogeneous. At network scale, the topology is complex resulting from mutual mechanical interactions as well as from major stresses. Geometric modeling of fractured networks combines in a non-standard way a large number of 2D fractures interconnected in the 3D space. Intricate local configurations of fracture intersections require original methods of geometric modeling and mesh generation. We have developed in 2016 a software package that automatically builds geometric models and surface meshes of random fracture networks. The results are highly promising and we now want to continue this research to further improve the element quality in complex configurations, take into account multiple size scales in large fracture networks (up to thousands of fractures), and compare several modeling strategies (mixed hybrid finite elements, projected grids, mortar elements).

4.3. Parallel meshing of surfaces defined by collections of connected regions

Participant: Patrick Laug [correspondant].

In CAD (computer aided design) environments, a surface is commonly modeled as a collection of connected regions represented by parametric mappings. For meshing such a composite surface, a parallelized indirect approach with dynamic load balancing can be used on a shared memory system. However, this methodology can be inefficient in practice because most existing CAD systems use memory caches that are only appropriate to a sequential process. As part of the sabbatical year of P. Laug at Polytechnique Montréal in 2014/2015, two solutions have been proposed, referred to as the Pirate approach and the Discrete approach. In the first approach, the Pirate library can be efficiently called in parallel since no caching is used for the storage or evaluation of geometric primitives. In the second approach, the CAD environment is replaced by internal procedures interpolating a discrete geometric support. In 2016, the dynamic load balancing has been analyzed and improved. Significant modifications to the Pirate library have been made, and new numerical tests on three different computers (4, 8 and 64 cores) have been carried out, now showing an almost linear scaling of the method in all cases.

4.4. Discrete CAD model for visualization and meshing

Participants: Patrick Laug [correspondant], Houman Borouchaki.

During the design of an object using a CAD (computer aided design) platform, the user can visualize the ongoing model at every moment. Visualization is based on a discrete representation of the model that coexists with the exact analytical representation of the object. Most CAD systems have this discrete representation available, and each of them applies its own construction methodology. We have developed in 2016 a method to build a discrete model for CAD surfaces (the model is quadtree-based and subdivided into quadrilaterals and triangles). The method presents two major particularities: most elements are aligned with iso-parametric curves and the accuracy of the surface approximation is controlled. In addition, we have proposed a new technique of surface mesh generation that is based on this discrete model. This approach has been implemented as a part of a surface mesher called ALIEN, and several examples have demonstrate the robustness and computational efficiency of the program, as well as the quality of the geometric support.

4.5. Visualization and modification of high-order curved meshes

Participants: Alexis Loyer, Dave Marcum, Adrien Loseille [correspondant].

During the partnership between Inria and Distene, a new visualization software has been designed. It address the typical operations that are required to quickly assess the newly algorithm developed in the team. In particular, interactive modifications of high-order curved mesh and hybrid meshes has been addressed.

4.6. Adaptation de maillages pour des écoulements visqueux en turbomachine

Participants: Frédéric Alauzet, Loïc Frazza, Adrien Loseille [correspondant].

4.6.1. Calcul.

Les prémices d'une adaptation pour les écoulements Navier-Stokes turbulents ont été testés sur des calculs de turbomachine. Pour ce faire nous avons tout d'abord traité les particularités liées aux calculs en turbomachine: - Les aubes présentent en général une périodicité par rotation et on ne simule donc qu'une période afin d'alléger les calculs. Il faut donc traiter cette périodicité de façon appropriée dans le code CFD et l'adaptation de maillage. - Afin de prendre en compte la rotation des pales sans employer de maillages mobiles et simulations instationnaires on peut se placer dans le référentiel tournant de l'aube en corrigeant les équations. - Les écoulements en turbomachine sont des écoulements clos, les conditions limites d'entrée et de sortie ont donc une influence très forte et peuvent de plus se trouver très près de la turbine afin de simuler la présence d'autres étages en amont ou aval. Des conditions limites bien précises ont donc été développées afin de traiter correctement ces effets.

4.6.2. Adaptation.

Pour l'adaptation de maillages deux particularités doivent être traitées ici, la périodicité du maillage et la couche limite turbulente.

En 2D, la couche limite turbulente est automatiquement adaptée avec la méthode metric orthogonal et la périodicité du maillage est garantie par un traitement spécial des frontières. Les estimateurs d'erreurs Navier-Stokes et RANS n'étant pas encore au point nous avons utilisé la Hessienne du Mach de l'écoulement comme senseur ce qui donne déjà des résultats satisfaisants.

En 3D la méthode metric orthogonal est beaucoup plus complexe à mettre en oeuvre et n'est pas encore au point. La couche limite a donc été exclue de l'adaptation, le maillage est adapté uniquement dans le volume en utilisant la Hessienne du Mach de l'écoulement comme senseur. La périodicité n'étant pas traitée non plus, les frontières périodiques restent inchangées ce qui garantie leur périodicité.

4.6.3. Norm-Oriented.

Dans le cadre de la théorie Norm-Oriented, afin de contrôler l'erreur implicite des schémas numériques, un correcteur a été développé et testé. Etant donné un maillage et la solution numérique obtenue avec, le résidu de cette solution projeté sur un maillage deux fois plus fin est accumulé sur le maillage initial. Ce défaut de résidu est utilisé comme terme source dans une seconde simulation plus courte. La nouvelle solution toujours sur le meme maillage est plus proche de la solution exacte et donne une bonne estimation de l'erreur.

4.7. Parallel mesh adaptation

Participants: Frédéric Alauzet, Adrien Loseille [correspondant].

We devise a strategy in order to generate large-size adapted anisotropic meshes $O(10^8 - 10^9)$ as required in many fields of application in scientific computing. We target moderate scale parallel computational resources as typically found in R&D units where the number of cores ranges in $O(10^2 - 10^3)$. Both distributed and shared memory architectures are handled. Our strategy is based on hierarchical domain splitting algorithm to remesh the partitions in parallel. Both the volume and the surface mesh are adapted simultaneously and the efficiency of the method is independent of the complexity of the geometry. The originality of the method relies on (i) a metric-based static load-balancing, (ii) dedicated hierarchical mesh partitioning techniques to (re)split the (complex) interfaces meshes, (iii) anisotropic Delaunay cavity to define the interface meshes, (iv) a fast, robust and generic sequential cavity-based mesh modification kernel, and (v) out-of-core storing of completing parts to reduce the memory footprint. We are able to generate (uniform, isotropic and anisotropic) meshes with more than 1 billion tetrahedra in less than 20 minutes on 120 cores.

4.8. Unsteady adjoint computation on dynamic meshes

Participants: Eléonore Gauci, Frédéric Alauzet [correspondant].

Adjoint formulations for unsteady problems are less common due to the extra complexity inherent in the numerical solution and storage but these methods are a great option in engineering because it takes more into account the cost function we want to minimize. Moreover the engineering applications involve moving bodies and this motion must be taken into account by the governing flow equations. We develop a model of unsteady adjoint solver on moving mesh problems. The derivation of the adjoint formulation based on the ALE form of the equations requires consideration of the dynamic meshes. Our model takes into account the DGCL.

4.9. Line solver for efficient stiff parse system resolution

Participants: Loïc Frazza, Frédéric Alauzet [correspondant].

Afin d'accélérer la résolution des problèmes raides, un line-solver á été développé. Cette méthode extrait tout d'abord des lignes dans le maillage du problème selon des critères géométriques ou physiques. Le problème peut alors être résolu exactement le long des ces lignes à moindre coût. Cette méthode est particulièrement bien adaptée aux cas où l'information se propage selon une direction privilégiée tels que les chocs, les couches limites ou les sillages. Ces cas sont généralement associés à des maillages très étirés ce qui conduit à des problèmes raides mais quasi-unidimensionnels. Ils peuvent donc être résolu efficacement par un line-solver, réduisant ainsi les temps de calculs tout en gagnant en robustesse.

4.10. Error estimate for high-order solution field

Participants: Olivier Coulaud, Adrien Loseille [correspondant].

Afin de produire des solveurs d'ordre élevé, et ainsi répondre aux exigences inhérentes à la résolution de problèmes physiques complexes, nous développons une méthode d'adaptation de maillage d'ordre élevé. Celle-ci est basée sur le contrôle par une métrique de l'erreur d'interpolation induite par le maillage du domaine. Plus précisément, pour une solution donnée, l'erreur d'interpolation d'ordre k est paramétrée par la forme différentielle $(k+1)^{\text{ième}}$ de cette solution, et le problème se réduit à trouver la plus grande ellipse incluse dans une ligne de niveau de cette différentielle. La méthode que nous avons mise au point théoriquement et numériquement est appelée "log-simplexe", et permet de produire des maillages adaptés d'ordre élevé dans un temps raisonnable, et ce en dimension 2 et 3. À l'occasion de l'International Meshing Roundtable 2016, ce travail a été présenté et publié. D'autres applications de cette méthode sont en cours d'exploitation, comme par exemple la génération de maillages adaptés courbes de surface, ou le couplage avec un solveur d'ordre élevé.

4.11. Méthode d'immersion de frontières pour la mécanique des fluides

Participants: Frédéric Alauzet [correspondant], Rémi Feuillet, Adrien Loseille.

Dans les méthodes de résolution classiques des problèmes d'interaction fluide-structure, il est usuel de représenter l'objet de manière exacte dans le maillage, c'est-à-dire avec des éléments conformes à l'objet : le maillage possède des triangles dont une arête correspond avec le bord de la géométrie immergée. Cette méthode quoique plus précise est très coûteuse en préprocessing. C'est dans ce cadre qu'est introduite la notion d'immersion de frontière (embedded geometry en anglais). Cette méthode consiste à représenter la géométrie de manière fictive. Le maillage de calcul n'est de fait plus nécessairement conforme à la géométrie de l'objet. Il s'agit donc de s'intéresser aux modifications nécessaires sur les méthodes classiques pour faire un calcul dans le cadre de l'immersion de frontières. Cela concerne les conditions aux limites et l'avancée en temps. On s'intéresse également à l'adaptation de maillage pour le cas de l'immersion. La finalité de tout ce travail est d'effectuer des calculs de coefficients aérodynamiques (portance, traînée) et de trouver des résultats du même ordre de précision que ceux en géométrie inscrite.

4.12. Boundary layer mesh generation

Participants: Frédéric Alauzet [correspondant], Adrien Loseille, Dave Marcum.

A closed advancing-layer method for generating high-aspect-ratio elements in the boundary layer (BL) region has been developed. This approach provides an answer to the mesh generation robustness issue as it starts from an existing valid mesh and always guarantuees a valid mesh in output. And, it handles very efficiently and naturally BL front collisions and it produces a natural smooth anisotropic blending between colliding layers. In addition, it provides a robust strategy to couple unstructured anisotropic mesh adaptation and high-aspect-ratio element pseudo-structured BL meshes. To this end, the mesh deformation is performed using the metric field associated with the given anisotropic meshes to maintain the adaptivity while inflating the BL. This approach utilizes a recently developed connectivity optimization based moving mesh strategy for deforming the volume mesh as the BL is inflated. In regards to the BL mesh generation, it features state-of-art capabilities, including, optimal normal evaluation, normal smoothing, blended BL termination, mixed-elements BL, varying growth rate, and BL imprinting on curved surfaces. Results for typical aerospace configurations are presented to assess the proposed strategy on both simple and complex geometries.

5. Bilateral Contracts and Grants with Industry

5.1. Bilateral Contracts with Industry

- The Boeing Company,
- Safran-Tech,
- Projet Rapid (DGA) avec Lemma.

6. Partnerships and Cooperations

6.1. European Initiatives

6.1.1. FP7 & H2020 Projects

- UMRIDA https://sites.google.com/a/numeca.be/umrida/

6.2. International Initiatives

6.2.1. Inria Associate Teams Not Involved in an Inria International Labs

6.2.1.1. AM2NS

Title: Advanced Meshing Methods for Numerical Simulations

International Partner (Institution - Laboratory - Researcher):

Mississippi State University (United States) - Center for Advanced Vehicular Systems - Computational Fluid Dynamics Dept. (CAVS-CFD) - Marcum David

Start year: 2017

See also: http://pages.saclay.inria.fr/frederic.alauzet/AssociateTeam_AM2NS/AT_am2ns.html

The purpose of the AM2NS Associate Team is to mutualize the knowledge of all teams in order to develop the next generation of meshing methods and their parallelization to address the new challenges in numerical simulations for industrial problems. The Associate Team is composed of four partners: Inria, Mississippi State University, The Boeing Company and Massachusetts Institute of Technology.

6.2.1.2. MODIS

Title: High-order discrete geometric modeling

International Partner (Institution - Laboratory - Researcher):

Polytechnique Montréal (Canada) - Computer Science - François Guibault

Start year: 2017

In the area of geometric modeling, major challenges are linked to the efficient visualization of CAD surfaces and to the generation of meshes adapted to numerical simulation. In this context, the conception of a discrete geometric model provides a simple and universal representation model, without the need for CAD. A first study has been carried out for the conception of a model of order 1 (one) defined by a “triangulation” composed of quadrilaterals and triangles. The advantage of this model of order 1 lies in its geometric simplicity. However, in the case of complex surfaces, it may require a very large number of elements, and besides it is not sufficiently rich to give certain essential characteristics like geometric curvatures. The main goal of this project is to extend this discrete model of order 1 to higher orders.

7. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] D. BARCHIESI, T. GROSGES.*Propagation of uncertainties and applications in numerical modeling: tutorial* , in "Journal of the Optical Society of America. A, Optics and image science", August 2017, vol. 34, n^{o} 9, p. 1602-1619 [*DOI :* 10.1364/JOSAA.34.001602], https://hal.inria.fr/hal-01575512.

[2] N. BARRAL, G. OLIVIER, F. ALAUZET.*Time-accurate anisotropic mesh adaptation for three-dimensional time-dependent problems with body-fitted moving geometries*, in "Journal of Computational Physics", February 2017, vol. 331, p. 157–187 [*DOI :* 10.1016/J.JCP.2016.11.029], https://hal.inria.fr/hal-01426156.

[3] A. CARABIAS, A. BELME, A. LOSEILLE, A. DERVIEUX.*Anisotropic Goal-oriented error analysis for a third-order accurate CENO Euler discretization*, in "International Journal for Numerical Methods in Fluids", 2017 [*DOI :* 10.1002/FLD.4423], http://hal.upmc.fr/hal-01579998.

International Conferences with Proceedings

[4] F. AMLANI, S. CHAILLAT, S. P. GROTH, A. LOSEILLE.*3D metric-based anisotropic mesh adaptation for the fast multipole accelerated boundary element method in acoustics*, in "WAVES 2017 - 13th International Conference on Mathematical and Numerical Aspects of Wave Propagation", Minneapolis, United States, May 2017, https://hal.archives-ouvertes.fr/hal-01598238.

[5] P. LAUG, H. BOROUCHAKI.*Geometric Meshing of Discrete Surfaces*, in "14th U.S. National Congress on Computational Mechanics (USNCCM) - Symposium on Trends in Unstructured Mesh Generation (STUMG)", Montreal, QC, Canada, July 2017, https://hal.inria.fr/hal-01591550.

[6] P. LAUG, G. PICHOT, J.-R. D. DREUZY.*Realistic geometric modeling of fracture networks*, in "8th International Conference on Adaptive Modeling and Simulation (ADMOS 2017) - Symposium "Mesh generation and mesh adaptativity: methods and applications"", Verbania, Italy, June 2017, https://hal.inria.fr/hal-01591579.

[7] A. LOSEILLE, R. FEUILLET.*Vizir: High-order mesh and solution visualization using OpenGL 4.0 graphic pipeline*, in "2018 - AIAA Aerospace Sciences Meeting, AIAA SciTech Forum", kissimmee, United States, January 2018, p. 1-13 [*DOI :* 10.2514/6.2018-1174], https://hal.inria.fr/hal-01686714.

Conferences without Proceedings

[8] L. GIRAUD-MOREAU, A. CHEROUAT, A. MERAT.*Numerical en Experimental study of the single point incremental forming of titanium sheet*, in "35th International Deep-Drawing Group Conference", Roissy-en-France, France, June 2017, https://hal.archives-ouvertes.fr/hal-01437756.

[9] A. LOSEILLE, A. LOYER, H. GUILLARD.*Coupling Grad-Shafranov equilibrium solvers and grid generation to study plasma confinement properties in nuclear fusion devices*, in "Coupled Problems 2017 - VII International Conference on Coupled Problems in Science and Engineering", Rhodes, Greece, June 2017, vol. 31, 190103, https://hal.inria.fr/hal-01644328.

Scientific Books (or Scientific Book chapters)

[10] A. LOSEILLE.*Unstructured Mesh Generation and Adaptation*, in "Handbook of Numerical Methods for Hyperbolic Problems — Applied and Modern Issues", R. ABGRALL, C.-W. SHU (editors), Elsevier , 2017, p. 263 - 302 [*DOI :* 10.1016/BS.HNA.2016.10.004], https://hal.inria.fr/hal-01438967.

Other Publications

[11] H. GUILLARD, J. LAKHLILI, A. LOSEILLE, A. LOYER, A. RATNANI.*Grid Generation for Fusion Applications*, October 2017, 1, EFTC 2017 - 17th European Fusion Theory Conference, Poster, https://hal.inria.fr/hal-01644309.

References in notes

[12] P.-G. CIARLET.*Basic Error Estimates for Elliptic Problems*, Ciarlet, P. G. and Lions, J. L., North Holland, 1991, vol. II.

[13] P.-J. FREY, P.-L. GEORGE.*Maillages. Applications aux éléments finis*, Hermès Science Publications, Paris, 1999.

Project-Team GECO

Geometric Control Design

IN PARTNERSHIP WITH:
CNRS

Ecole Polytechnique

RESEARCH CENTER
Saclay - Île-de-France

THEME
Optimization and control of dynamic systems

Table of contents

Project-Team GECO

Creation of the Team: 2011 May 01, updated into Project-Team: 2013 January 01, end of the Project-Team: 2017 December 31

Keywords:

Computer Science and Digital Science:

A1.5. - Complex systems
A5.3. - Image processing and analysis
A6.1. - Mathematical Modeling
A6.4.1. - Deterministic control
A6.4.3. - Observability and Controlability
A6.4.4. - Stability and Stabilization

Other Research Topics and Application Domains:

B1.2.1. - Understanding and simulation of the brain and the nervous system
B2.6. - Biological and medical imaging
B9.4.2. - Mathematics
B9.4.3. - Physics

1. Personnel

Research Scientists

Mario Sigalotti [Team leader, Inria, Researcher, until Jun 2017, HDR]
Ugo Boscain [CNRS, Senior Researcher, until Jun 2017, HDR]

PhD Students

Nicolas Augier [Ecole polytechnique, until Jun 2017]
Mathieu Kohli [Ecole polytechnique, until Jun 2017]
Jakub Orlowski [Centrale-Supélec, until Jun 2017]
Ludovic Sacchelli [Ecole polytechnique, until Jun 2017]

Post-Doctoral Fellows

Francesco Boarotto [Ecole polytechnique, until Jun 2017]
Valentina Franceschi [Inria, until Jun 2017]

Administrative Assistant

Tiffany Caristan [Inria]

2. Overall Objectives

2.1. Overall Objectives

Motion planning is not only a crucial issue in control theory, but also a widespread task of all sort of human activities. The aim of the project-team is to study the various aspects preceding and framing *motion planning*: accessibility analysis (determining which configurations are attainable), criteria to make choice among possible trajectories, trajectory tracking (fixing a possibly unfeasible trajectory and following it as closely as required), performance analysis (determining the cost of a tracking strategy), design of implementable algorithms, robustness of a control strategy with respect to computationally motivated discretizations, etc. The viewpoint that we adopt comes from geometric control: our main interest is in qualitative and intrinsic properties and our focus is on trajectories (either individual ones or families of them).

The main application domain of GECO is *quantum control*. The importance of designing efficient transfers between different atomic or molecular levels in atomic and molecular physics is due to its applications to photochemistry (control by laser pulses of chemical reactions), nuclear magnetic resonance (control by a magnetic field of spin dynamics) and, on a more distant time horizon, the strategic domain of quantum computing.

A second application area concerns the control interpretation of phenomena appearing in *neurophysiology*. It studies the modeling of the mechanisms supervising some biomechanics actions or sensorial reactions such as image reconstruction by the primary visual cortex, eyes movement and body motion. All these problems can be seen as motion planning tasks accomplished by the brain.

As a third applicative domain we propose a system dynamics approach to *switched systems*. Switched systems are characterized by the interaction of continuous dynamics (physical system) and discrete/logical components. They provide a popular modeling framework for heterogeneous aspects issuing from automotive and transportation industry, energy management and factory automation.

3. Research Program

3.1. Geometric control theory

The main research topic of the project-team is **geometric control**, with a special focus on **control design**. The application areas that we target are control of quantum mechanical systems, neurogeometry and switched systems.

Geometric control theory provides a viewpoint and several tools, issued in particular from differential geometry, to tackle typical questions arising in the control framework: controllability, observability, stabilization, optimal control... [9], [43] The geometric control approach is particularly well suited for systems involving nonlinear and nonholonomic phenomena. We recall that nonholonomicity refers to the property of a velocity constraint that is not equivalent to a state constraint.

The expression **control design** refers here to all phases of the construction of a control law, in a mainly open-loop perspective: modeling, controllability analysis, output tracking, motion planning, simultaneous control algorithms, tracking algorithms, performance comparisons for control and tracking algorithms, simulation and implementation.

We recall that

- **controllability** denotes the property of a system for which any two states can be connected by a trajectory corresponding to an admissible control law ;
- **output tracking** refers to a control strategy aiming at keeping the value of some functions of the state arbitrarily close to a prescribed time-dependent profile. A typical example is **configuration tracking** for a mechanical system, in which the controls act as forces and one prescribes the position variables along the trajectory, while the evolution of the momenta is free. One can think for instance at the lateral movement of a car-like vehicle: even if such a movement is unfeasible, it can be tracked with arbitrary precision by applying a suitable control strategy;
- **motion planning** is the expression usually denoting the algorithmic strategy for selecting one control law steering the system from a given initial state to an attainable final one;
- **simultaneous control** concerns algorithms that aim at driving the system from two different initial conditions, with the same control law and over the same time interval, towards two given final states (one can think, for instance, at some control action on a fluid whose goal is to steer simultaneously two floating bodies.) Clearly, the study of which pairs (or n-uples) of states can be simultaneously connected thanks to an admissible control requires an additional controllability analysis with respect to the plain controllability mentioned above.

At the core of control design is then the notion of motion planning. Among the motion planning methods, a preeminent role is played by those based on the Lie algebra associated with the control system ([63], [50], [56]), those exploiting the possible flatness of the system ([37]) and those based on the continuation method ([75]). Optimal control is clearly another method for choosing a control law connecting two states, although it generally introduces new computational and theoretical difficulties.

Control systems with special structure, which are very important for applications are those for which the controls appear linearly. When the controls are not bounded, this means that the admissible velocities form a distribution in the tangent bundle to the state manifold. If the distribution is equipped with a smoothly varying norm (representing a cost of the control), the resulting geometrical structure is called *sub-Riemannian*. Sub-Riemannian geometry thus appears as the underlying geometry of the nonholonomic control systems, playing the same role as Euclidean geometry for linear systems. As such, its study is fundamental for control design. Moreover its importance goes far beyond control theory and is an active field of research both in differential geometry ([62]), geometric measure theory ([38], [13]) and hypoelliptic operator theory ([25]).

Other important classes of control systems are those modeling mechanical systems. The dynamics are naturally defined on the tangent or cotangent bundle of the configuration manifold, they have Lagrangian or Hamiltonian structure, and the controls act as forces. When the controls appear linearly, the resulting model can be seen somehow as a second-order sub-Riemannian structure (see [30]).

The control design topics presented above naturally extend to the case of distributed parameter control systems. The geometric approach to control systems governed by partial differential equations is a novel subject with great potential. It could complement purely analytical and numerical approaches, thanks to its more dynamical, qualitative and intrinsic point of view. An interesting example of this approach is the paper [10] about the controllability of Navier–Stokes equation by low forcing modes.

4. Application Domains

4.1. Quantum control

The issue of designing efficient transfers between different atomic or molecular levels is crucial in atomic and molecular physics, in particular because of its importance in those fields such as photochemistry (control by laser pulses of chemical reactions), nuclear magnetic resonance (NMR, control by a magnetic field of spin dynamics) and, on a more distant time horizon, the strategic domain of quantum computing. This last application explicitly relies on the design of quantum gates, each of them being, in essence, an open loop control law devoted to a prescribed simultaneous control action. NMR is one of the most promising techniques for the implementation of a quantum computer.

Physically, the control action is realized by exciting the quantum system by means of one or several external fields, being them magnetic or electric fields. The resulting control problem has attracted increasing attention, especially among quantum physicists and chemists (see, for instance, [68], [73]). The rapid evolution of the domain is driven by a multitude of experiments getting more and more precise and complex (see the recent review [29]). Control strategies have been proposed and implemented, both on numerical simulations and on physical systems, but there is still a large gap to fill before getting a complete picture of the control properties of quantum systems. Control techniques should necessarily be innovative, in order to take into account the physical peculiarities of the model and the specific experimental constraints.

The area where the picture got clearer is given by finite dimensional linear closed models.

- **Finite dimensional** refers to the dimension of the space of wave functions, and, accordingly, to the finite number of energy levels.
- **Linear** means that the evolution of the system for a fixed (constant in time) value of the control is determined by a linear vector field.
- **Closed** refers to the fact that the systems are assumed to be totally disconnected from the environment, resulting in the conservation of the norm of the wave function.

The resulting model is well suited for describing spin systems and also arises naturally when infinite dimensional quantum systems of the type discussed below are replaced by their finite dimensional Galerkin approximations. Without seeking exhaustiveness, let us mention some of the issues that have been tackled for finite dimensional linear closed quantum systems:

- controllability [11],
- bounds on the controllability time [7],
- STIRAP processes [78],
- simultaneous control [51],
- optimal control ([47], [20], [31]),
- numerical simulations [57].

Several of these results use suitable transformations or approximations (for instance the so-called rotating wave) to reformulate the finite-dimensional Schrödinger equation as a sub-Riemannian system. Open systems have also been the object of an intensive research activity (see, for instance, [12], [48], [69], [26]).

In the case where the state space is infinite dimensional, some optimal control results are known (see, for instance, [16], [27], [44], [17]). The controllability issue is less understood than in the finite dimensional setting, but several advances should be mentioned. First of all, it is known that one cannot expect exact controllability on the whole Hilbert sphere [77]. Moreover, it has been shown that a relevant model, the quantum oscillator, is not even approximately controllable [70], [60]. These negative results have been more recently completed by positive ones. In [18], [19] Beauchard and Coron obtained the first positive controllability result for a quantum particle in a 1D potential well. The result is highly nontrivial and is based on Coron's return method (see [33]). Exact controllability is proven to hold among regular enough wave functions. In particular, exact controllability among eigenfunctions of the uncontrolled Schrödinger operator can be achieved. Other important approximate controllability results have then been proved using Lyapunov methods [59], [64], [45]. While [59] studies a controlled Schrödinger equation in $\mathbb{R}$ for which the uncontrolled Schrödinger operator has mixed spectrum, [64], [45] deal mainly with general discrete-spectrum Schrödinger operators.

In all the positive results recalled in the previous paragraph, the quantum system is steered by a single external field. Different techniques can be applied in the case of two or more external fields, leading to additional controllability results [36], [23].

The picture is even less clear for nonlinear models, such as Gross–Pitaevski and Hartree–Fock equations. The obstructions to exact controllability, similar to the ones mentioned in the linear case, have been discussed in [42]. Optimal control approaches have also been considered [15], [28]. A comprehensive controllability analysis of such models is probably a long way away.

4.2. Neurophysiology

At the interface between neurosciences, mathematics, automatics and humanoid robotics, an entire new approach to neurophysiology is emerging. It arouses a strong interest in the four communities and its development requires a joint effort and the sharing of complementary tools.

A family of extremely interesting problems concerns the understanding of the mechanisms supervising some sensorial reactions or biomechanics actions such as image reconstruction by the primary visual cortex, eyes movement and body motion.

In order to study these phenomena, a promising approach consists in identifying the motion planning problems undertaken by the brain, through the analysis of the strategies that it applies when challenged by external inputs. The role of control is that of a language allowing to read and model neurological phenomena. The control algorithms would shed new light on the brain's geometric perception (the so-called neurogeometry [66]) and on the functional organization of the motor pathways.

- A challenging problem is that of the understanding of the mechanisms which are responsible for the process of image reconstruction in the primary visual cortex V1.

The visual cortex areas composing V1 are notable for their complex spatial organization and their functional diversity. Understanding and describing their architecture requires sophisticated modeling tools. At the same time, the structure of the natural and artificial images used in visual psychophysics can be fully disclosed only using rather deep geometric concepts. The word “geometry" refers here to the internal geometry of the functional architecture of visual cortex areas (not to the geometry of the Euclidean external space). Differential geometry and analysis both play a fundamental role in the description of the structural characteristics of visual perception.

A model of human perception based on a simplified description of the visual cortex V1, involving geometric objects typical of control theory and sub-Riemannian geometry, has been first proposed by Petitot ([67]) and then modified by Citti and Sarti ([32]). The model is based on experimental observations, and in particular on the fundamental work by Hubel and Wiesel [41] who received the Nobel prize in 1981.

In this model, neurons of V1 are grouped into orientation columns, each of them being sensitive to visual stimuli arriving at a given point of the retina and oriented along a given direction. The retina is modeled by the real plane, while the directions at a given point are modeled by the projective line. The fiber bundle having as base the real plane and as fiber the projective line is called the *bundle of directions of the plane*.

From the neurological point of view, orientation columns are in turn grouped into hypercolumns, each of them sensitive to stimuli arriving at a given point, oriented along any direction. In the same hypercolumn, relative to a point of the plane, we also find neurons that are sensitive to other stimuli properties, such as colors. Therefore, in this model the visual cortex treats an image not as a planar object, but as a set of points in the bundle of directions of the plane. The reconstruction is then realized by minimizing the energy necessary to activate orientation columns among those which are not activated directly by the image. This gives rise to a sub-Riemannian problem on the bundle of directions of the plane.

- Another class of challenging problems concern the functional organization of the motor pathways.

  The interest in establishing a model of the motor pathways, at the same time mathematically rigorous and biologically plausible, comes from the possible spillovers in robotics and neurophysiology. It could help to design better control strategies for robots and artificial limbs, yielding smoother and more progressive movements. Another underlying relevant societal goal (clearly beyond our domain of expertise) is to clarify the mechanisms of certain debilitating troubles such as cerebellar disease, chorea and Parkinson’s disease.

  A key issue in order to establish a model of the motor pathways is to determine the criteria underlying the brain’s choices. For instance, for the problem of human locomotion (see [14]), identifying such criteria would be crucial to understand the neural pathways implicated in the generation of locomotion trajectories.

  A nowadays widely accepted paradigm is that, among all possible movements, the accomplished ones satisfy suitable optimality criteria (see [76] for a review). One is then led to study an inverse optimal control problem: starting from a database of experimentally recorded movements, identify a cost function such that the corresponding optimal solutions are compatible with the observed behaviors.

  Different methods have been taken into account in the literature to tackle this kind of problems, for instance in the linear quadratic case [46] or for Markov processes [65]. However all these methods have been conceived for very specific systems and they are not suitable in the general case. Two approaches are possible to overcome this difficulty. The direct approach consists in choosing a cost function among a class of functions naturally adapted to the dynamics (such as energy functions) and to compare the solutions of the corresponding optimal control problem to the experimental data. In particular one needs to compute, numerically or analytically, the optimal trajectories and to choose suitable criteria (quantitative and qualitative) for the comparison with observed trajectories. The inverse approach consists in deriving the cost function from the qualitative analysis of the data.

4.3. Switched systems

Switched systems form a subclass of hybrid systems, which themselves constitute a key growth area in automation and communication technologies with a broad range of applications. Existing and emerging areas include automotive and transportation industry, energy management and factory automation. The notion of hybrid systems provides a framework adapted to the description of the heterogeneous aspects related to the interaction of continuous dynamics (physical system) and discrete/logical components.

The characterizing feature of switched systems is the collective aspect of the dynamics. A typical question is that of stability, in which one wants to determine whether a dynamical system whose evolution is influenced by a time-dependent signal is uniformly stable with respect to all signals in a fixed class ([53]).

The theory of finite-dimensional hybrid and switched systems has been the subject of intensive research in the last decade and a large number of diverse and challenging problems such as stabilizability, observability, optimal control and synchronization have been investigated (see for instance [74], [54]).

The question of stability, in particular, because of its relevance for applications, has spurred a rich literature. Important contributions concern the notion of common Lyapunov function: when there exists a Lyapunov function that decays along all possible modes of the system (that is, for every possible constant value of the signal), then the system is uniformly asymptotically stable. Conversely, if the system is stable uniformly with respect to all signals switching in an arbitrary way, then a common Lyapunov function exists [55]. In the *linear* finite-dimensional case, the existence of a common Lyapunov function is actually equivalent to the global uniform exponential stability of the system [61] and, provided that the admissible modes are finitely many, the Lyapunov function can be taken polyhedral or polynomial [21], [22], [34]. A special role in the switched control literature has been played by common quadratic Lyapunov functions, since their existence can be tested rather efficiently (see [35] and references therein). Algebraic approaches to prove the stability of switched systems under arbitrary switching, not relying on Lyapunov techniques, have been proposed in [52], [8].

Other interesting issues concerning the stability of switched systems arise when, instead of considering arbitrary switching, one restricts the class of admissible signals, by imposing, for instance, a dwell time constraint [40].

Another rich area of research concerns discrete-time switched systems, where new intriguing phenomena appear, preventing the algebraic characterization of stability even for small dimensions of the state space [49]. It is known that, in this context, stability cannot be tested on periodic signals alone [24].

Finally, let us mention that little is known about infinite-dimensional switched system, with the exception of some results on uniform asymptotic stability ([58], [71], [72]) and some recent papers on optimal control ([39], [79]).

5. Highlights of the Year

5.1. Highlights of the Year

GECO has ended in June 2017, after being evaluated earlier in the year. A new team, including all former members of GECO, has started in July 2017 in the Inria Paris center. Its name is CAGE, for *Control And GEometry*.

6. New Software and Platforms

6.1. IRHD

Image Reconstruction via Hypoelliptic Diffusion

FUNCTIONAL DESCRIPTION: IRHD is a software for reconstruction of corrupted and damaged images. One of the main features of the algorithm on which the software is based is that it does not require any information about the location and character of the corrupted places. Another important advantage is that this method is massively parallelizable, this allows to work with sufficiently large images. Theoretical background of the presented method is based on the model of geometry of vision due to Petitot, Citti and Sarti. The main step is numerical solution of the equation of 3D hypoelliptic diffusion. IRHD is based on Fortran.

- Contact: Mario Sigalotti

7. New Results

7.1. New results

Let us list some the new results in sub-Riemannian geometry and hypoelliptic diffusion obtained by GECO's members.

- On a sub-Riemannian manifold we define two type of Laplacians. The macroscopic Laplacian Δ_ω, as the divergence of the horizontal gradient, once a volume ω is fixed, and the microscopic Laplacian, as the operator associated with a geodesic random walk. In [1] we consider a general class of random walks, where all sub-Riemannian geodesics are taken in account. This operator depends only on the choice of a complement c to the sub-Riemannian distribution, and is denoted L_c. We address the problem of equivalence of the two operators. This problem is interesting since, on equiregular sub-Riemannian manifolds, there is always an intrinsic volume (e.g. Popp's one P) but not a canonical choice of complement. The result depends heavily on the type of structure under investigation:
  - On contact structures, for every volume ω, there exists a unique complement c such that $\Delta_\omega = L_c$.
  - On Carnot groups, if H is the Haar volume, then there always exists a complement c such that $\Delta_H = L_c$. However this complement is not unique in general.
  - For quasi-contact structures, in general, $\Delta_P = L_c$ for any choice of c. In particular, L_c is not symmetric w.r.t. Popp's measure. This is surprising especially in dimension 4 where, in a suitable sense, Δ_P is the unique intrinsic macroscopic Laplacian.

  A crucial notion that we introduce here is the N-intrinsic volume, i.e. a volume that depends only on the set of parameters of the nilpotent approximation. When the nilpotent approximation does not depend on the point, a N-intrinsic volume is unique up to a scaling by a constant and the corresponding N-intrinsic sub-Laplacian is unique. This is what happens for dimension smaller or equal than 4, and in particular in the 4-dimensional quasi-contact structure mentioned above.
- In sub-Riemannian geometry the coefficients of the Jacobi equation define curvature-like invariants. We show in [4] that these coefficients can be interpreted as the curvature of a canonical Ehresmann connection associated to the metric, first introduced by Zelenko and Li. We show why this connection is naturally nonlinear, and we discuss some of its properties.
- In [6] we study the cut locus of the free, step two Carnot groups $\mathbb{G}_k$ with k generators, equipped with their left-invariant Carnot-Carathéodory metric. In particular, we disprove the conjectures on the shape of the cut loci proposed by O. Myasnichenko, by exhibiting sets of cut points $C_k \subset \mathbb{G}_k$ which, for $k \geq 4$, are strictly larger than conjectured ones. While the latter were, respectively, smooth semi-algebraic sets of codimension $\Theta(k^2)$ and semi-algebraic sets of codimension $\Theta(k)$, the sets C_k are semi-algebraic and have codimension 2, yielding the best possible lower bound valid for all k on the size of the cut locus of $\mathbb{G}_k$. Furthermore, we study the relation of the cut locus with the so-called abnormal set. Finally, and as a straightforward consequence of our results, we derive an explicit lower bound for the small time heat kernel asymptotics at the points of C_k. The question whether C_k coincides with the cut locus for $k \geq 4$ remains open.

New results on complex systems with hybrid or switched components are the following.

- In [2] we address the exponential stability of a system of transport equations with intermittent damping on a network of $N \geq 2$ circles intersecting at a single point O. The N equations are coupled through a linear mixing of their values at O, described by a matrix M. The activity of the intermittent damping is determined by persistently exciting signals, all belonging to a fixed class. The main result is that, under suitable hypotheses on M and on the rationality of the ratios between the lengths of the circles, such a system is exponentially stable, uniformly with respect to the persistently exciting signals. The proof relies on an explicit formula for the solutions of this system, which allows one to track down the effects of the intermittent damping.
- In [3] we study the relative controllability of linear difference equations with multiple delays in the state by using a suitable formula for the solutions of such systems in terms of their initial conditions, their control inputs, and some matrix-valued coefficients obtained recursively from the matrices defining the system. Thanks to such formula, we characterize relative controllability in time T in terms of an algebraic property of the matrix-valued coefficients, which reduces to the usual Kalman controllability criterion in the case of a single delay. Relative controllability is studied for solutions in the set of all functions and in the function spaces L^p and C^k. We also compare the relative controllability of the system for different delays in terms of their rational dependence structure, proving that relative controllability for some delays implies relative controllability for all delays that are "less rationally dependent" than the original ones, in a sense that we make precise. Finally, we provide an upper bound on the minimal controllability time for a system depending only on its dimension and on its largest delay.

Finally, a new contribution has been proposed in the domain of the control of quantum systems. More precisely, in [5] we consider the bilinear Schrödinger equation with discrete-spectrum drift. We show, for $n \in \mathbb{N}$ arbitrary, exact controllability in projections on the first n given eigenstates. The controllability result relies on a generic controllability hypothesis on some associated finite-dimensional approximations. The method is based on Lie-algebraic control techniques applied to the finite-dimensional approximations coupled with classical topological arguments issuing from degree theory.

8. Partnerships and Cooperations

8.1. Regional Initiatives

- Starting from the end of 2015, we have been funded by PGMO (Gaspard Monge Program for Optimisation and operational research) through a grant on Geometric Optimal Control. The grant is coordinated by Mario Sigalotti.

8.2. National Initiatives

8.2.1. ANR

The ANR SRGI starts at the end of 2015, for a duration of four years. GECO is one of one of the partners of the ANR. The national coordinator is Emmanuel Trélat (UPMC) and the local one Ugo Boscain.

SRGI deals with sub-Riemannian geometry, hypoelliptic diffusion and geometric control.

8.2.2. Other initiatives

Ugo Boscain and Mario Sigalotti are members of the project DISQUO of the program Inphyniti of the CNRS. Coordinator: Thomas Chambrion (Nancy).

8.3. European Initiatives

8.3.1. FP7 & H2020 Projects

Program: ERC Proof of Concept

Project acronym: ARTIV1

Project title: An artificial visual cortex for image processing

Duration: From April 2017 to September 2018.

Coordinator: Ugo Boscain

Abstract: The ERC starting grant GECOMETHODS, on which this POC is based, tackled problems of diffusion equations via geometric control methods. One of the most striking achievements of the project has been the development of an algorithm of image reconstruction based mainly on non-isotropic diffusion. This algorithm is bio-mimetic in the sense that it replicates the way in which the primary visual cortex V1 of mammals processes the signals arriving from the eyes. It has performances that are at the state of the art in image processing. These results together with others obtained in the ERC project show that image processing algorithms based on the functional architecture of V1 can go very far. However, the exceptional performances of the primary visual cortex V1 rely not only on the particular algorithm used, but also on the fact that such algorithm ?runs? on a dedicated hardware having the following features: 1. an exceptional level of parallelism; 2. connections that are well adapted to transmit information in a non-isotropic way as it is required by the algorithms of image reconstruction and recognition. The idea of this POC is to create a dedicated hardware (called ARTIV1) emulating the functional architecture of V1 and hence having on one hand a huge degree of parallelism and on the other hand connections among the CPUs that reflect the non-isotropic structure of the visual cortex V1. Such a hardware that we plan to build as an integrated circuit with an industrial partner will be a veritable artificial visual cortex. It will be fully programmable and it will be able to perform many biomimetic image processing tasks that we expect to be exceptionally performant. ARTIV1 will come to the marked accompanied by some dedicated software for image reconstruction and image recognition. However we expect that other applications will be developed by customers, as for instance softwares for optical flow estimation or for sound processing.

8.4. International Initiatives

8.4.1. Informal International Partners

SISSA (Scuola Internazionale Superiore di Studi Avanzati), Trieste, Italy.

Sector of Functional Analysis and Applications, Geometric Control group. Coordinator: Andrei A. Agrachev.

We collaborate with the Geometric Control group at SISSA mainly on subjects related with sub-Riemannian geometry. Thanks partly to our collaboration, SISSA has established an official research partnership with École Polytechnique.

8.5. International Research Visitors

8.5.1. Visits of International Scientists

- Andrei Agrachev (SISSA, Italy) has been visiting the GECO team for one year, ending in June 2017.

8.5.1.1. Internships

Gontran Lance has made an internship in GECO, under the supervision of Mario Sigalotti and Emmanuel Trélat on the turnpike phenomenon in the orbital transfer problem.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

Ugo Boscain was organizer of the conference “Mathematical Control Theory, with a special session in honor of Gianna Stefani”, Porquerolles, 27–30 June.

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards

- Ugo Boscain is Associate Editor of SIAM Journal of Control and Optimization
- Ugo Boscain is Managing Editor of Journal of Dynamical and Control Systems
- Mario Sigalotti is Associate Editor of Journal of Dynamical and Control Systems
- Mario Sigalotti is Associate Editor of ESAIM Control, Optimisation and Calculus of Variations
- Ugo Boscain is Associate Editor of Mathematical Control and Related Fields
- Ugo Boscain is Associate editor of Analysis and Geometry in Metric Spaces

9.1.3. Invited Talks

- Mario Sigalotti gave an invited talk at the conference “Mathematical Control Theory, with a special session in honor of Gianna Stefani”, Porquerolles, France, June 2017.
- Mario Sigalotti gave an invited talk at groupe de travail “Contrôle”, Laboratoire Jacques-Louis Lions, Paris, France, January 2017.

9.1.4. Research Administration

- Mario Sigalotti is member of the IFAC technical committee “Distributed Parameter Systems”.
- Mario Sigalotti was member of the steering committee of the *Institut pour le Contrôle et la Décision* of the Idex Paris-Saclay up to June 2017.

9.2. Teaching - Supervision - Juries

9.2.1. Supervision

- PhD in progress: Ludovic Sacchelli, “Sub-Riemannian geometry, hypoelliptic operators, geometry of vision", started in September 2015, supervisors: Ugo Boscain, Mario Sigalotti.
- PhD in progress: Nicolas Augier, “Contrôle adiabatique des systèmes quantiques", started in September 2016, supervisors: Ugo Boscain, Mario Sigalotti.
- PhD in progress: Mathieu Kohli, “Volume and curvature in sub-Riemannian geometry", started in September 2016, supervisors: Davide Barilari, Ugo Boscain.
- PhD in progress: Jakub Orłowski, “Modeling and steering brain oscillations based on in vivo optogenetics data", started in September 2016, supervisors: Antoine Chaillet, Alain Destexhe, and Mario Sigalotti.

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] U. Boscain, R. Neel, L. Rizzi. *Intrinsic random walks and sub-Laplacians in sub-Riemannian geometry*, in "Advances in Mathematics", July 2017, https://arxiv.org/abs/1503.00725 [*DOI :* 10.1016/j.aim.2017.04.024], https://hal.archives-ouvertes.fr/hal-01122735.

[2] Y. CHITOUR, G. MAZANTI, M. SIGALOTTI.*Persistently damped transport on a network of circles*, in "Transactions of the American Mathematical Society", June 2017, vol. 369, n^{o} 6, p. 3841-3881, https://arxiv.org/abs/1406.0731 [*DOI :* 10.1090/TRAN/6778], https://hal.inria.fr/hal-00999743.

[3] G. MAZANTI.*Relative controllability of linear difference equations*, in "SIAM Journal on Control and Optimization", 2017, vol. 55, n^{o} 5, p. 3132–3153, https://arxiv.org/abs/1604.08663 [*DOI :* 10.1137/16M1073157], https://hal.archives-ouvertes.fr/hal-01309166.

Other Publications

[4] D. BARILARI, L. RIZZI.*On Jacobi fields and canonical connection in sub-Riemannian geometry*, March 2017, https://arxiv.org/abs/1506.01827 - Final version, to appear on Archivum Mathematicum, https://hal.archives-ouvertes.fr/hal-01160902.

[5] M. CAPONIGRO, M. SIGALOTTI.*Exact controllability in projections of the bilinear Schrödinger equation*, 2017, working paper or preprint, https://hal.inria.fr/hal-01509971.

[6] L. RIZZI, U. SERRES.*On the cut locus of free, step two Carnot groups*, January 2017, https://arxiv.org/abs/1610.01596 - 13 pages. To appear on Proceedings of the AMS, https://hal.archives-ouvertes.fr/hal-01377408.

References in notes

[7] A. A. AGRACHEV, T. CHAMBRION.*An estimation of the controllability time for single-input systems on compact Lie groups*, in "ESAIM Control Optim. Calc. Var.", 2006, vol. 12, n^{o} 3, p. 409–441.

[8] A. A. AGRACHEV, D. LIBERZON.*Lie-algebraic stability criteria for switched systems*, in "SIAM J. Control Optim.", 2001, vol. 40, n^{o} 1, p. 253–269, http://dx.doi.org/10.1137/S0363012999365704.

[9] A. A. AGRACHEV, Y. L. SACHKOV.*Control theory from the geometric viewpoint*, Encyclopaedia of Mathematical Sciences, Springer-Verlag, Berlin, 2004, vol. 87, xiv+412, Control Theory and Optimization, II.

[10] A. A. AGRACHEV, A. V. SARYCHEV.*Navier-Stokes equations: controllability by means of low modes forcing*, in "J. Math. Fluid Mech.", 2005, vol. 7, n^{o} 1, p. 108–152, http://dx.doi.org/10.1007/s00021-004-0110-1.

[11] F. ALBERTINI, D. D'ALESSANDRO.*Notions of controllability for bilinear multilevel quantum systems*, in "IEEE Trans. Automat. Control", 2003, vol. 48, n^{o} 8, p. 1399–1403.

[12] C. ALTAFINI.*Controllability properties for finite dimensional quantum Markovian master equations*, in "J. Math. Phys.", 2003, vol. 44, n^{o} 6, p. 2357–2372.

[13] L. AMBROSIO, P. TILLI.*Topics on analysis in metric spaces*, Oxford Lecture Series in Mathematics and its Applications, Oxford University Press, Oxford, 2004, vol. 25, viii+133.

[14] G. ARECHAVALETA, J.-P. LAUMOND, H. HICHEUR, A. BERTHOZ.*An optimality principle governing human locomotion*, in "IEEE Trans. on Robotics", 2008, vol. 24, n^{o} 1.

[15] L. BAUDOUIN.*A bilinear optimal control problem applied to a time dependent Hartree-Fock equation coupled with classical nuclear dynamics*, in "Port. Math. (N.S.)", 2006, vol. 63, n^{o} 3, p. 293–325.

[16] L. BAUDOUIN, O. KAVIAN, J.-P. PUEL.*Regularity for a Schrödinger equation with singular potentials and application to bilinear optimal control*, in "J. Differential Equations", 2005, vol. 216, n^{o} 1, p. 188–222.

[17] L. BAUDOUIN, J. SALOMON.*Constructive solution of a bilinear optimal control problem for a Schrödinger equation*, in "Systems Control Lett.", 2008, vol. 57, n^{o} 6, p. 453–464, http://dx.doi.org/10.1016/j.sysconle.2007.11.002.

[18] K. BEAUCHARD.*Local controllability of a 1-D Schrödinger equation*, in "J. Math. Pures Appl. (9)", 2005, vol. 84, n^{o} 7, p. 851–956.

[19] K. BEAUCHARD, J.-M. CORON.*Controllability of a quantum particle in a moving potential well*, in "J. Funct. Anal.", 2006, vol. 232, n^{o} 2, p. 328–389.

[20] M. BELHADJ, J. SALOMON, G. TURINICI.*A stable toolkit method in quantum control*, in "J. Phys. A", 2008, vol. 41, n^{o} 36, 362001, 10, http://dx.doi.org/10.1088/1751-8113/41/36/362001.

[21] F. BLANCHINI.*Nonquadratic Lyapunov functions for robust control*, in "Automatica J. IFAC", 1995, vol. 31, n^{o} 3, p. 451–461, http://dx.doi.org/10.1016/0005-1098(94)00133-4.

[22] F. BLANCHINI, S. MIANI.*A new class of universal Lyapunov functions for the control of uncertain linear systems*, in "IEEE Trans. Automat. Control", 1999, vol. 44, n^{o} 3, p. 641–647, http://dx.doi.org/10.1109/9.751368.

[23] A. M. BLOCH, R. W. BROCKETT, C. RANGAN.*Finite Controllability of Infinite-Dimensional Quantum Systems*, in "IEEE Trans. Automat. Control", 2010.

[24] V. D. BLONDEL, J. THEYS, A. A. VLADIMIROV.*An elementary counterexample to the finiteness conjecture*, in "SIAM J. Matrix Anal. Appl.", 2003, vol. 24, n^{o} 4, p. 963–970, http://dx.doi.org/10.1137/S0895479801397846.

[25] A. BONFIGLIOLI, E. LANCONELLI, F. UGUZZONI.*Stratified Lie groups and potential theory for their sub-Laplacians*, Springer Monographs in Mathematics, Springer, Berlin, 2007, xxvi+800.

[26] B. BONNARD, D. SUGNY.*Time-minimal control of dissipative two-level quantum systems: the integrable case*, in "SIAM J. Control Optim.", 2009, vol. 48, n^{o} 3, p. 1289–1308, http://dx.doi.org/10.1137/080717043.

[27] A. BORZÌ, E. DECKER.*Analysis of a leap-frog pseudospectral scheme for the Schrödinger equation*, in "J. Comput. Appl. Math.", 2006, vol. 193, n^{o} 1, p. 65–88.

[28] A. BORZÌ, U. HOHENESTER.*Multigrid optimization schemes for solving Bose-Einstein condensate control problems*, in "SIAM J. Sci. Comput.", 2008, vol. 30, n^{o} 1, p. 441–462, http://dx.doi.org/10.1137/070686135.

[29] C. BRIF, R. CHAKRABARTI, H. RABITZ.*Control of quantum phenomena: Past, present, and future*, Advances in Chemical Physics, S. A. Rice (ed), Wiley, New York, 2010.

[30] F. BULLO, A. D. LEWIS.*Geometric control of mechanical systems*, Texts in Applied Mathematics, Springer-Verlag, New York, 2005, vol. 49, xxiv+726.

[31] R. CABRERA, H. RABITZ.*The landscape of quantum transitions driven by single-qubit unitary transformations with implications for entanglement*, in "J. Phys. A", 2009, vol. 42, n^{o} 27, 275303, 9, http://dx.doi.org/10.1088/1751-8113/42/27/275303.

[32] G. CITTI, A. SARTI.*A cortical based model of perceptual completion in the roto-translation space*, in "J. Math. Imaging Vision", 2006, vol. 24, n^{o} 3, p. 307–326, http://dx.doi.org/10.1007/s10851-005-3630-2.

[33] J.-M. CORON.*Control and nonlinearity*, Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 2007, vol. 136, xiv+426.

[34] W. P. DAYAWANSA, C. F. MARTIN.*A converse Lyapunov theorem for a class of dynamical systems which undergo switching*, in "IEEE Trans. Automat. Control", 1999, vol. 44, n^{o} 4, p. 751–760, http://dx.doi.org/10.1109/9.754812.

[35] L. EL GHAOUI, S.-I. NICULESCU.*Robust decision problems in engineering: a linear matrix inequality approach*, in "Advances in linear matrix inequality methods in control", Philadelphia, PA, Adv. Des. Control, SIAM, 2000, vol. 2, p. 3–37.

[36] S. ERVEDOZA, J.-P. PUEL.*Approximate controllability for a system of Schrödinger equations modeling a single trapped ion*, in "Ann. Inst. H. Poincaré Anal. Non Linéaire", 2009, vol. 26, p. 2111–2136.

[37] M. FLIESS, J. LÉVINE, P. MARTIN, P. ROUCHON.*Flatness and defect of non-linear systems: introductory theory and examples*, in "Internat. J. Control", 1995, vol. 61, n^{o} 6, p. 1327–1361, http://dx.doi.org/10.1080/00207179508921959.

[38] B. FRANCHI, R. SERAPIONI, F. SERRA CASSANO.*Regular hypersurfaces, intrinsic perimeter and implicit function theorem in Carnot groups*, in "Comm. Anal. Geom.", 2003, vol. 11, n^{o} 5, p. 909–944.

[39] M. GUGAT.*Optimal switching boundary control of a string to rest in finite time*, in "ZAMM Z. Angew. Math. Mech.", 2008, vol. 88, n^{o} 4, p. 283–305.

[40] J. HESPANHA, S. MORSE.*Stability of switched systems with average dwell-time*, in "Proceedings of the 38th IEEE Conference on Decision and Control, CDC 1999, Phoenix, AZ, USA", 1999, p. 2655–2660.

[41] D. HUBEL, T. WIESEL.*Brain and Visual Perception: The Story of a 25-Year Collaboration*, Oxford University Press, Oxford, 2004.

[42] R. ILLNER, H. LANGE, H. TEISMANN.*Limitations on the control of Schrödinger equations*, in "ESAIM Control Optim. Calc. Var.", 2006, vol. 12, n^{o} 4, p. 615–635, http://dx.doi.org/10.1051/cocv:2006014.

[43] A. ISIDORI.*Nonlinear control systems*, Communications and Control Engineering Series, Second, Springer-Verlag, Berlin, 1989, xii+479, An introduction.

[44] K. ITO, K. KUNISCH.*Optimal bilinear control of an abstract Schrödinger equation*, in "SIAM J. Control Optim.", 2007, vol. 46, n^{o} 1, p. 274–287.

[45] K. ITO, K. KUNISCH.*Asymptotic properties of feedback solutions for a class of quantum control problems*, in "SIAM J. Control Optim.", 2009, vol. 48, n^{o} 4, p. 2323–2343, http://dx.doi.org/10.1137/080720784.

[46] R. KALMAN.*When is a linear control system optimal?*, in "ASME Transactions, Journal of Basic Engineering", 1964, vol. 86, p. 51–60.

[47] N. KHANEJA, S. J. GLASER, R. W. BROCKETT.*Sub-Riemannian geometry and time optimal control of three spin systems: quantum gates and coherence transfer*, in "Phys. Rev. A (3)", 2002, vol. 65, n^{o} 3, part A, 032301, 11.

[48] N. KHANEJA, B. LUY, S. J. GLASER.*Boundary of quantum evolution under decoherence*, in "Proc. Natl. Acad. Sci. USA", 2003, vol. 100, n^{o} 23, p. 13162–13166, http://dx.doi.org/10.1073/pnas.2134111100.

[49] V. S. KOZYAKIN.*Algebraic unsolvability of a problem on the absolute stability of desynchronized systems*, in "Avtomat. i Telemekh.", 1990, p. 41–47.

[50] G. LAFFERRIERE, H. J. SUSSMANN.*A differential geometry approach to motion planning*, in "Nonholonomic Motion Planning (Z. Li and J. F. Canny, editors)", Kluwer Academic Publishers, 1993, p. 235-270.

[51] J.-S. LI, N. KHANEJA.*Ensemble control of Bloch equations*, in "IEEE Trans. Automat. Control", 2009, vol. 54, n^{o} 3, p. 528–536, http://dx.doi.org/10.1109/TAC.2009.2012983.

[52] D. LIBERZON, J. P. HESPANHA, A. S. MORSE.*Stability of switched systems: a Lie-algebraic condition*, in "Systems Control Lett.", 1999, vol. 37, n^{o} 3, p. 117–122, http://dx.doi.org/10.1016/S0167-6911(99)00012-2.

[53] D. LIBERZON.*Switching in systems and control*, Systems & Control: Foundations & Applications, Birkhäuser Boston Inc., Boston, MA, 2003, xiv+233.

[54] H. LIN, P. J. ANTSAKLIS.*Stability and stabilizability of switched linear systems: a survey of recent results*, in "IEEE Trans. Automat. Control", 2009, vol. 54, n^{o} 2, p. 308–322, http://dx.doi.org/10.1109/TAC.2008.2012009.

[55] Y. LIN, E. D. SONTAG, Y. WANG.*A smooth converse Lyapunov theorem for robust stability*, in "SIAM J. Control Optim.", 1996, vol. 34, n^{o} 1, p. 124–160, http://dx.doi.org/10.1137/S0363012993259981.

[56] W. LIU.*Averaging theorems for highly oscillatory differential equations and iterated Lie brackets*, in "SIAM J. Control Optim.", 1997, vol. 35, n^{o} 6, p. 1989–2020, http://dx.doi.org/10.1137/S0363012994268667.

[57] Y. MADAY, J. SALOMON, G. TURINICI.*Monotonic parareal control for quantum systems*, in "SIAM J. Numer. Anal.", 2007, vol. 45, n^{o} 6, p. 2468–2482, http://dx.doi.org/10.1137/050647086.

[58] A. N. MICHEL, Y. SUN, A. P. MOLCHANOV.*Stability analysis of discountinuous dynamical systems determined by semigroups*, in "IEEE Trans. Automat. Control", 2005, vol. 50, n^{o} 9, p. 1277–1290, http://dx.doi.org/10.1109/TAC.2005.854582.

[59] M. MIRRAHIMI.*Lyapunov control of a particle in a finite quantum potential well*, in "Proceedings of the 45th IEEE Conference on Decision and Control", 2006.

[60] M. MIRRAHIMI, P. ROUCHON.*Controllability of quantum harmonic oscillators*, in "IEEE Trans. Automat. Control", 2004, vol. 49, n^{o} 5, p. 745–747.

[61] A. P. MOLCHANOV, Y. S. PYATNITSKIY.*Criteria of asymptotic stability of differential and difference inclusions encountered in control theory*, in "Systems Control Lett.", 1989, vol. 13, n^{o} 1, p. 59–64, http://dx.doi.org/10.1016/0167-6911(89)90021-2.

[62] R. MONTGOMERY.*A tour of subriemannian geometries, their geodesics and applications*, Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 2002, vol. 91, xx+259.

[63] R. M. MURRAY, S. S. SASTRY.*Nonholonomic motion planning: steering using sinusoids*, in "IEEE Trans. Automat. Control", 1993, vol. 38, n^{o} 5, p. 700–716, http://dx.doi.org/10.1109/9.277235.

[64] V. NERSESYAN.*Growth of Sobolev norms and controllability of the Schrödinger equation*, in "Comm. Math. Phys.", 2009, vol. 290, n^{o} 1, p. 371–387.

[65] A. Y. NG, S. RUSSELL.*Algorithms for Inverse Reinforcement Learning*, in "Proc. 17th International Conf. on Machine Learning", 2000, p. 663–670.

[66] J. PETITOT.*Neurogéomètrie de la vision. Modèles mathématiques et physiques des architectures fonctionnelles*, Les Éditions de l'École Polythechnique, 2008.

[67] J. PETITOT, Y. TONDUT.*Vers une neurogéométrie. Fibrations corticales, structures de contact et contours subjectifs modaux*, in "Math. Inform. Sci. Humaines", 1999, n^{o} 145, p. 5–101.

[68] H. RABITZ, H. DE VIVIE-RIEDLE, R. MOTZKUS, K. KOMPA.*Wither the future of controlling quantum phenomena?*, in "SCIENCE", 2000, vol. 288, p. 824–828.

[69] D. ROSSINI, T. CALARCO, V. GIOVANNETTI, S. MONTANGERO, R. FAZIO.*Decoherence by engineered quantum baths*, in "J. Phys. A", 2007, vol. 40, n^{o} 28, p. 8033–8040, http://dx.doi.org/10.1088/1751-8113/40/28/S12.

[70] P. ROUCHON.*Control of a quantum particle in a moving potential well*, in "Lagrangian and Hamiltonian methods for nonlinear control 2003", Laxenburg, IFAC, 2003, p. 287–290.

[71] A. SASANE.*Stability of switching infinite-dimensional systems*, in "Automatica J. IFAC", 2005, vol. 41, n^{o} 1, p. 75–78, http://dx.doi.org/10.1016/j.automatica.2004.07.013.

[72] A. SAURABH, M. H. FALK, M. B. ALEXANDRE.*Stability analysis of linear hyperbolic systems with switching parameters and boundary conditions*, in "Proceedings of the 47th IEEE Conference on Decision and Control, CDC 2008, December 9-11, 2008, Cancún, Mexico", 2008, p. 2081–2086.

[73] M. SHAPIRO, P. BRUMER.*Principles of the Quantum Control of Molecular Processes*, Principles of the Quantum Control of Molecular Processes, pp. 250. Wiley-VCH, February 2003.

[74] R. SHORTEN, F. WIRTH, O. MASON, K. WULFF, C. KING.*Stability criteria for switched and hybrid systems*, in "SIAM Rev.", 2007, vol. 49, n^{o} 4, p. 545–592, http://dx.doi.org/10.1137/05063516X.

[75] H. J. SUSSMANN.*A continuation method for nonholonomic path finding*, in "Proceedings of the 32th IEEE Conference on Decision and Control, CDC 1993, Piscataway, NJ, USA", 1993, p. 2718–2723.

[76] E. TODOROV.*12*, in "Optimal control theory", Bayesian Brain: Probabilistic Approaches to Neural Coding, Doya K (ed), 2006, p. 269–298.

[77] G. TURINICI.*On the controllability of bilinear quantum systems*, in "Mathematical models and methods for ab initio Quantum Chemistry", M. DEFRANCESCHI, C. LE BRIS (editors), Lecture Notes in Chemistry, Springer, 2000, vol. 74.

[78] L. YATSENKO, S. GUÉRIN, H. JAUSLIN.*Topology of adiabatic passage*, in "Phys. Rev. A", 2002, vol. 65, 043407, 7.

[79] E. ZUAZUA.*Switching controls*, in "Journal of the European Mathematical Society", 2011, vol. 13, n^{o} 1, p. 85–117.

Project-Team GRACE

Geometry, arithmetic, algorithms, codes and encryption

IN COLLABORATION WITH: Laboratoire d’informatique de l’école polytechnique (LIX)

IN PARTNERSHIP WITH:
CNRS

Ecole Polytechnique

RESEARCH CENTER
Saclay - Île-de-France

THEME
Algorithmics, Computer Algebra and Cryptology

Table of contents

Project-Team GRACE

Creation of the Team: 2012 January 01, updated into Project-Team: 2013 July 01

Keywords:

Computer Science and Digital Science:

A4.2. - Correcting codes
A4.3. - Cryptography
A4.3.1. - Public key cryptography
A4.3.3. - Cryptographic protocols
A4.8. - Privacy-enhancing technologies
A8.4. - Computer Algebra
A8.5. - Number theory

Other Research Topics and Application Domains:

B9.4.2. - Mathematics
B9.8. - Privacy

1. Personnel

Research Scientists

Daniel Augot [Team leader, Inria, Senior Researcher, HDR]
Alain Couvreur [Inria, Researcher]
William George [Laboratoire d'informatique de l'école polytechnique (LIX), Starting Research Position]
Matthieu Rambaud [Inria, Researcher, from Oct 2017]
Benjamin Smith [Inria, Researcher]

Faculty Members

Luca de Feo [Univ de Versailles Saint-Quentin-en-Yvelines, Associate Professor]
Francoise Levy-Dit-Vehel [École Nationale Supérieure de Techniques Avancées, Associate Professor, HDR]
François Morain [Ecole polytechnique, Professor, HDR]

External Collaborators

Christian Berghoff [Bonn International Graduate School - BIGS, until Feb 2017]
Pierre Karpman [Centrum Wiskunde and Informatica, Amsterdam, until Nov 2017]
David Kohel [Aix–Marseille Université,from Jul 2017]
Philippe Lebacque [Univ de Franche-Comté]

Technical Staff

Nicholas Coxon [Inria]

PhD Students

Elise Barelli [Inria]
Hanna-Mae Bisserier [Institut de recherche technologique System X]
Nicolas Duhamel [Ecole Polytechnique, until Jul 2017]
Hussein Khazaie [Inria, from Oct 2017]
Julien Lavauzelle [Ecole polytechnique]
Edouard Rousseau [Institut Telecom ex GET Groupe des Ecoles des Télécommunications , from Oct 2017]

Visiting Scientist

Elisabeth Malmskog [Colorado College, USA, from Nov 2017]

Administrative Assistant

Jessica Gameiro [Inria]

2. Overall Objectives

2.1. Scientific foundations

GRACE has two broad application domains—cryptography and coding theory—linked by a common foundation in algorithmic number theory and the geometry of algebraic curves. In our research, which combines theoretical work with practical software development, we use algebraic curves to *create better cryptosystems*, to *provide better security assessments* for cryptographic key sizes, and to *build the best error-correcting codes*.

Coding and cryptography deal (in different ways) with securing communication systems for high-level applications. In our research, the two domains are linked by the computational issues related to algebraic curves (over various fields) and arithmetic rings. These fundamental number-theoretic algorithms, at the crossroads of a rich area of mathematics and computer science, have already proven their relevance in public key cryptography, with industrial successes including the RSA cryptosystem and elliptic curve cryptography. It is less well-known that the same branches of mathematics can be used to build very good codes for error correction. While coding theory has traditionally had an electrical engineering flavour, recent developments in computer science have shed new light on coding theory, leading to new applications more central to computer science.

3. Research Program

3.1. Algorithmic Number Theory

Algorithmic Number Theory is concerned with replacing special cases with general algorithms to solve problems in number theory. In the Grace project, it appears in three main threads:

- fundamental algorithms for integers and polynomials (including primality and factorization);
- algorithms for finite fields (including discrete logarithms); and
- algorithms for algebraic curves.

Clearly, we use computer algebra in many ways. Research in cryptology has motivated a renewed interest in Algorithmic Number Theory in recent decades—but the fundamental problems still exist *per se*. Indeed, while algorithmic number theory application in cryptanalysis is epitomized by applying factorization to breaking RSA public key, many other problems, are relevant to various area of computer science. Roughly speaking, the problems of the cryptological world are of bounded size, whereas Algorithmic Number Theory is also concerned with asymptotic results.

3.2. Arithmetic Geometry: Curves and their Jacobians

Theme: Arithmetic Geometry: Curves and their Jacobians *Arithmetic Geometry* is the meeting point of algebraic geometry and number theory: that is, the study of geometric objects defined over arithmetic number systems (such as the integers and finite fields). The fundamental objects for our applications in both coding theory and cryptology are curves and their Jacobians over finite fields.

An algebraic *plane curve*$\mathcal{X}$ over a field $\mathbf{K}$ is defined by an equation

$$\mathcal{X} : F_{\mathcal{X}}(x, y) = 0 \quad \text{where } F_{\mathcal{X}} \in \mathbf{K}[x, y].$$

(Not every curve is planar—we may have more variables, and more defining equations—but from an algorithmic point of view, we can always reduce to the plane setting.) The *genus* $g_{\mathcal{X}}$ of $\mathcal{X}$ is a non-negative integer classifying the essential geometric complexity of $\mathcal{X}$; it depends on the degree of $F_{\mathcal{X}}$ and on the number of singularities of $\mathcal{X}$. The curve $\mathcal{X}$ is associated in a functorial way with an algebraic group $J_{\mathcal{X}}$, called the *Jacobian* of $\mathcal{X}$. The group $J_{\mathcal{X}}$ has a geometric structure: its elements correspond to points on a $g_{\mathcal{X}}$-dimensional projective algebraic group variety. Typically, we do not compute with the equations defining this projective variety: there are too many of them, in too many variables, for this to be convenient. Instead, we use fast algorithms based on the representation in terms of classes of formal sums of points on $\mathcal{X}$.

The simplest curves with nontrivial Jacobians are curves of genus 1, known as *elliptic curves*; they are typically defined by equations of the form $y^2 = x^3 + Ax + B$. Elliptic curves are particularly important given their central role in public-key cryptography over the past two decades. Curves of higher genus are important in both cryptography and coding theory.

3.3. Curve-Based cryptology

Theme: Curve-Based Cryptology

Jacobians of curves are excellent candidates for cryptographic groups when constructing efficient instances of public-key cryptosystems. Diffie–Hellman key exchange is an instructive example.

Suppose Alice and Bob want to establish a secure communication channel. Essentially, this means establishing a common secret *key*, which they will then use for encryption and decryption. Some decades ago, they would have exchanged this key in person, or through some trusted intermediary; in the modern, networked world, this is typically impossible, and in any case completely unscalable. Alice and Bob may be anonymous parties who want to do e-business, for example, in which case they cannot securely meet, and they have no way to be sure of each other's identities. Diffie–Hellman key exchange solves this problem. First, Alice and Bob publicly agree on a cryptographic group G with a generator P (of order N); then Alice secretly chooses an integer a from $[1..N]$, and sends aP to Bob. In the meantime, Bob secretly chooses an integer b from $[1..N]$, and sends bP to Alice. Alice then computes $a(bP)$, while Bob computes $b(aP)$; both have now computed abP, which becomes their shared secret key. The security of this key depends on the difficulty of computing abP given P, aP, and bP; this is the Computational Diffie–Hellman Problem (CDHP). In practice, the CDHP corresponds to the Discrete Logarithm Problem (DLP), which is to determine a given P and aP.

This simple protocol has been in use, with only minor modifications, since the 1970s. The challenge is to create examples of groups G with a relatively compact representation and an efficiently computable group law, and such that the DLP in G is hard (ideally approaching the exponential difficulty of the DLP in an abstract group). The Pohlig–Hellman reduction shows that the DLP in G is essentially only as hard as the DLP in its largest prime-order subgroup. We therefore look for compact and efficient groups of prime order.

The classic example of a group suitable for the Diffie–Hellman protocol is the multiplicative group of a finite field $\mathbf{F}_q$. There are two problems that render its usage somewhat less than ideal. First, it has too much structure: we have a subexponential Index Calculus attack on the DLP in this group, so while it is very hard, the DLP falls a long way short of the exponential difficulty of the DLP in an abstract group. Second, there is only one such group for each q: its subgroup treillis depends only on the factorization of $q - 1$, and requiring $q - 1$ to have a large prime factor eliminates many convenient choices of q.

This is where Jacobians of algebraic curves come into their own. First, elliptic curves and Jacobians of genus 2 curves do not have a subexponential index calculus algorithm: in particular, from the point of view of the DLP, a generic elliptic curve is currently *as strong as* a generic group of the same size. Second, they provide some diversity: we have many degrees of freedom in choosing curves over a fixed $\mathbf{F}_q$, with a consequent diversity of possible cryptographic group orders. Furthermore, an attack which leaves one curve vulnerable may not necessarily apply to other curves. Third, viewing a Jacobian as a geometric object rather than a pure group allows us to take advantage of a number of special features of Jacobians. These features include efficiently computable pairings, geometric transformations for optimised group laws, and the availability of efficiently computable non-integer endomorphisms for accelerated encryption and decryption.

3.4. Algebraic Coding Theory

Theme: Coding theory

Coding Theory studies originated with the idea of using redundancy in messages to protect against noise and errors. The last decade of the 20th century has seen the success of so-called iterative decoding methods, which enable us to get very close to the Shannon capacity. The capacity of a given channel is the best achievable transmission rate for reliable transmission. The consensus in the community is that this capacity is more easily reached with these iterative and probabilistic methods than with algebraic codes (such as Reed–Solomon codes).

However, algebraic coding is useful in settings other than the Shannon context. Indeed, the Shannon setting is a random case setting, and promises only a vanishing error probability. In contrast, the algebraic Hamming approach is a worst case approach: under combinatorial restrictions on the noise, the noise can be adversarial, with strictly zero errors.

These considerations are renewed by the topic of list decoding after the breakthrough of Guruswami and Sudan at the end of the nineties. List decoding relaxes the uniqueness requirement of decoding, allowing a small list of candidates to be returned instead of a single codeword. List decoding can reach a capacity close to the Shannon capacity, with zero failure, with small lists, in the adversarial case. The method of Guruswami and Sudan enabled list decoding of most of the main algebraic codes: Reed–Solomon codes and Algebraic–Geometry (AG) codes and new related constructions “capacity-achieving list decodable codes”. These results open the way to applications again adversarial channels, which correspond to worst case settings in the classical computer science language.

Another avenue of our studies is AG codes over various geometric objects. Although Reed–Solomon codes are the best possible codes for a given alphabet, they are very limited in their length, which cannot exceed the size of the alphabet. AG codes circumvent this limitation, using the theory of algebraic curves over finite fields to construct long codes over a fixed alphabet. The striking result of Tsfasman–Vladut–Zink showed that codes better than random codes can be built this way, for medium to large alphabets. Disregarding the asymptotic aspects and considering only finite length, AG codes can be used either for longer codes with the same alphabet, or for codes with the same length with a smaller alphabet (and thus faster underlying arithmetic).

From a broader point of view, wherever Reed–Solomon codes are used, we can substitute AG codes with some benefits: either beating random constructions, or beating Reed–Solomon codes which are of bounded length for a given alphabet.

Another area of Algebraic Coding Theory with which we are more recently concerned is the one of Locally Decodable Codes. After having been first theoretically introduced, those codes now begin to find practical applications, most notably in cloud-based remote storage systems.

4. Highlights of the Year

4.1. Highlights of the Year

4.1.1. Presentation at Inria@SiliconValley

D. Augot made a presentation at a one day workshop “Blockchain Technology for Cybersecurity and Social Impact” at Berkeley’s CITRIS https://project.inria.fr/siliconvalley/bis2017-day1-conference-blockchain

4.1.2. Workshop on Coding theory and Cryptography (WCC)

D. Augot was co-chair of the Program Commitee of WCC 2017 (St Petersburg, Russia).

4.1.3. NIST Call for post quantum cryptography

In the context of NIST’s call for post quantum cryptography:

https://csrc.nist.gov/Projects/Post-Quantum-Cryptography
members of the team participated to two sumbissions:

- A. Couvreur and E. Barelli participated to the submission of **BIG QUAKE** proposal [19]:
  https://bigquake.inria.fr/
- L. De Feo participated to the submission of **SIKE** proposal:
  https://rwc.iacr.org/2018//Slides/Longa.pdf

5. New Software and Platforms

5.1. ACTIS

Algorithmic Coding Theory in Sage
FUNCTIONAL DESCRIPTION: The aim of this project is to vastly improve the state of the error correcting library in Sage. The existing library does not present a good and usable API, and the provided algorithms are very basic, irrelevant, and outdated. We thus have two directions for improvement: renewing the APIs to make them actually usable by researchers, and incorporating efficient programs for decoding, like J. Nielsen's CodingLib, which contains many new algorithms.

- Partner: Technical University Denmark
- Contact: Daniel Augot

5.2. DECODING

KEYWORD: Algebraic decoding
FUNCTIONAL DESCRIPTION: Decoding is a standalone C library. Its primary goal is to implement Guruswami–Sudan list decoding-related algorithms, as efficiently as possible. Its secondary goal is to give an efficient tool for the implementation of decoding algorithms (not necessarily list decoding algorithms) and their benchmarking.

- Participant: Guillaume Quintin
- Contact: Daniel Augot

5.3. Fast Compact Diffie-Hellman

KEYWORD: Cryptography
FUNCTIONAL DESCRIPTION: A competitive, high-speed, open implementation of the Diffie–Hellman protocol, targeting the 128-bit security level on Intel platforms. This download contains Magma files that demonstrate how to compute scalar multiplications on the x-line of an elliptic curve using endomorphisms. This accompanies the EuroCrypt 2014 paper by Costello, Hisil and Smith, the full version of which can be found here: http://eprint.iacr.org/2013/692 . The corresponding SUPERCOP-compatible crypto_dh application can be downloaded from http://hhisil.yasar.edu.tr/files/hisil20140318compact.tar.gz .

- Participant: Benjamin Smith
- Contact: Benjamin Smith
- URL: http://research.microsoft.com/en-us/downloads/ef32422a-af38-4c83-a033-a7aafbc1db55/

5.4. CADO-NFS

Crible Algébrique: Distribution, Optimisation - Number Field Sieve
KEYWORDS: Cryptography - Number theory

FUNCTIONAL DESCRIPTION: CADO-NFS is a complete implementation in C/C++ of the Number Field Sieve (NFS) algorithm for factoring integers and computing discrete logarithms in finite fields. It consists in various programs corresponding to all the phases of the algorithm, and a general script that runs them, possibly in parallel over a network of computers.

- Participants: Pierrick Gaudry, Emmanuel Thomé and Paul Zimmermann
- Contact: Emmanuel Thomé
- URL: http://cado-nfs.gforge.inria.fr/

6. New Results

6.1. qDSA: Compact signatures for IoT

B. Smith and Joost Renes (Radboud University, NL) developed **qDSA**, a new digital signature scheme targeting constrained devices, typically microcontrollers with extremely limited memory. An article describing qDSA was presented at ASIACRYPT 2017, and a reference implementation software package has been placed into the public domain.

6.2. PIR based on transversal designs

J. Lavauzelle presented a construction of Private Information Retrieval (PIR) protocols from combinatorial structures called transversal designs. The construction features low computation and low storage overhead for the servers. For some instances, adequate communication between servers and user is achieved. The PIR scheme also generalizes to colluding servers. The construction has been presented during WCC 2017 [17], and in a poster session in the Munich Workshop in Coding and Applications.

6.3. On the security of compact McEliece keys

E. Barelli presented at WCC 2017 (Workshop on Coding and Cryptography, St Petersburg, Russia) her recent results on the analysis of McEliece scheme based on alternant codes with a non trivial automorphism group [16]. These codes were suggested for public key encryption since, compared to codes with trivial automorphism group, they could provide shorter keys.

If the security with respect to generic decoding attacks is almost unchanged when considering codes with non trivial automorphisms, E. Barelli proved that the security with respect to key recovery attacks is highly reduced since, it reduces to recover the structure of the subcode of fixed elements by the automorphism group.

6.4. Two-points codes on the generalized Giuletti Korchmaros curve

In a collaboration with Peter Beelen, Mrinmoy Datta, Vincent Neiger and Johan Rosenkilde (DTU Copenhagen), E. Barelli obtained improved lower bounds for the minimum distance of some algebraic geometry codes from Giulietti Korchmaros curves [20].

6.5. Towards a function field version of Freiman's theorem

In a collaboration with Christine Bachoc and Gilles Zémor (University of Bordeaux), A. Couvreur obtained a characterisation of subspaces S of a function field F over an algebraically closed field satisfiying

$$\dim S^2 = 2 \dim S$$

where S^2 denotes the space spanned by all the products of two elements of S. They obtained the following result [18]:

Theorem. *Let F be a function field over an algebraically closed field, and S be a finite dimensional subspace of F which spans F as an algebra and such that*

$$\dim S^2 = 2\dim S.$$

Then F is a function field of transcendence degree 1 and

- *either F has genus 1 and S is a Riemann Roch space*
- *or F has genus 0 and S is a subspace of codimension 1 in a Riemann Roch space.*

6.6. BIG QUAKE

In the context of NIST's call for post quantum cryptosystems:

https://csrc.nist.gov/Projects/Post-Quantum-Cryptography

A. Couvreur and E. Barelli participated to the submission BIG QUAKE [19] (BInary Goppa QUAsi–cyclic Key Encapsulation). The proposal consists in a public key encryption scheme (with a conversion to a Key Encapsulation Mechanism) using binary quasi–cyclic Goppa codes.

The details on the proposal are on the following website.

https://bigquake.inria.fr/

6.7. Discrete Logarithm computations in finite fields with the NFS algorithm

The best discrete logarithm record computations in prime fields and large characteristic finite fields are obtained with Number Field Sieve algorithm (NFS) at the moment.

6.7.1. Computing discrete logarithms in $GF(p^6)$

A. Guillevic, L. Grémy, F. Morain and E. Thomé (from CARAMBA EPC in LORIA) computed a discrete log on a curve of embedding degree 6 and cryptographic size. This clearly showed that curves with small embedding degrees are indeed weak. The article [23] was presented by L. Grémy during the SAC 2017 conference in Ottawa.

6.7.2. Identity management on Bitcoin's blockchain

D. Augot and W. George in collaboration with Hervé Chabanne (Safran Identity and Security, ex Morpho, now Idemia) designed two schemes to allow users to authenticate using so-called anonymous credentials, issues by an identity provider. We used Brands anonymous credentials with selective disclosure each time, first for a finely tuned, user managed, identity scheme [12], second for a more classical high throughput scheme [13], inspired by CONIKS https://coniks.cs.princeton.edu.

6.7.3. Law and Blokchain smart contracts

D. Augot, with Célia Zolynski, is co-advising Hanna-Mae Bisserier, a PhD student law, on the impact of blockchains on legal systems. The PhD is in law, and D. Augot only gives scientific and technological explanations, while the direction of the thesis is done by Célia Zolynski.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

- **NOKIA BELL LABS**
  - New PhD student H. Khazaie is funded by ADR with NOKIA BELL LABS. The PhD topic is the security of distributed storage systems.
  - Post doctoral researcher N. Coxon is funded by ADR with NOKIA BELL LABS. The post doc topic is an information theoretically secure private information retrieval scheme.
- **SAFRAN Identity and Security (Ex Morpho and now Idemia)**
  - Post doctoral researcher W. George is funded by Idemia to design an identity management scheme based on Bitcoin's blockchain.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

MANTA (accepted July 2015, starting March 2016): "Curves, surfaces, codes and cryptography". This project deals with applications of coding theory error correcting codes to in cryptography, multi-party computation, and complexity theory, using advanced topics in algebraic geometry and number theory. The kickoff was a one week-retreat in Dordogne (20 participants), and we had another four day meeting in Saclay in November 17. See http://anr-manta.inria.fr/.

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

8.2.1.1. PQCRYPTO

Title: Post-quantum cryptography for long-term security

Programm: H2020

Duration: March 2015 - March 2018

Coordinator: TECHNISCHE UNIVERSITEIT EINDHOVEN

Partners:

Academia Sinica (Taiwan)

Bundesdruckerei (Germany)

Danmarks Tekniske Universitet (Denmark)

Katholieke Universiteit Leuven (Belgium)

Nxp Semiconductors Belgium Nv (Belgium)

Ruhr-Universitaet Bochum (Germany)

Stichting Katholieke Universiteit (Netherlands)

Coding Theory and Cryptology group, Technische Universiteit Eindhoven (Netherlands)

Technische Universitaet Darmstadt (Germany)

University of Haifa (Israel)

Inria contact: Nicolas Sendrier

Online security depends on a very few underlying cryptographic algorithms. Public-key algorithms are particularly crucial since they provide digital signatures and establish secure communication. Essentially all applications today are based on RSA or on the discrete-logarithm problem in finite fields or on elliptic curves. Cryptographers optimize parameter choices and implementation details for these systems and build protocols on top of these systems; cryptanalysts fine-tune attacks and establish exact security levels for these systems.

It might seem that having three systems offers enough variation, but these systems are all broken as soon as large quantum computers are built. The EU and governments around the world are investing heavily in building quantum computers; society needs to be prepared for the consequences, including cryptanalytic attacks accelerated by these computers. Long-term confidential documents such as patient health-care records and state secrets have to guarantee security for many years, but information encrypted today using RSA or elliptic curves and stored until quantum computers are available will then be as easy to decipher.

PQCRYPTO will allow users to switch to post-quantum cryptography: cryptographic systems that are not merely secure for today but that will also remain secure long-term against attacks by quantum computers. PQCRYPTO will design a portfolio of high-security post-quantum public-key systems, and will improve the speed of these systems, with reference implementations.

Our team is engaged in WP3.3 “advanced applications for the cloud”. We envision to focus essentially on secure multiparty computation, essentially the information theoretically secure constructions, who are naturally secure against a quantum computer invoked on classical queries. We will study whether these protocols still resist quantum queries. This work sub package started March 2015, and is dealt with by D. Augot.

8.3. International Initiatives

8.3.1. Inria International Partners

8.3.1.1. Informal International Partners

B. Smith has continued our successful informal partnership with the cryptography research group at Radboud University, Nijmegen (NL). 2017 has seen visits from researchers in both directions, and the production of the **qDSA** signature scheme package.

8.4. International Research Visitors

8.4.1. Visits of International Scientists

Beth Malmskog (Colorado College) visited the team from November 27 to December 1 2017 and gave a talk on locally recoverable codes based on fibre products of algebraic curves.

8.4.2. Visits to International Teams

8.4.2.1. Research Stays Abroad

B. Smith was an invited researcher in the Computer Science department at CINVESTAV (Mexico City, Mexico) for the month of August 2017, hosted by Professor Francisco Rodríguez Henríquez.

J. Lavauzelle visited Incidence Geometry team at Gent University (Belgium) for the month of April 2017, hosted by Professor Leo Storme.

E. Barelli visited the COMPUTE team in the DTU University at Lyngby (Danemark) during one month in february-march 2017, hosted by Professor Peter Beelen.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. Member of the Organizing Committees

B. Smith was a member of the organizing committee and Short Talk Chair for IEEE EuroS&P 2017 (Paris, April 2017)

9.1.2. Scientific Events Selection

9.1.2.1. Chair of Conference Program Committees

- D. Augot was co-chair of Workshop on Coding and Cryptography (WCC) 2017 at St Petersburg (Russia).

9.1.2.2. Member of the Conference Program Committees

B. Smith: Latincrypt 2017, ECC (International Workshop on Elliptic Curve Cryptography) 2017.

D. Augot and A. Couvreur : Fifth Code-based Cryptography Workshop 2017, Tenerife, Spain.

A. Couvreur: WCC 2017 (Workshop on Coding and Cryptography 2017, St Petersbug, Russia).

A. Couvreur : AGC^{2}T 2017 (Arithmetic Geometry Cryptography and Coding Theory 2017, Marseille, France).

D. Augot: International Conference on Mathematical Aspects of Computer and Information Sciences https://macis2017.sba-research.org/

9.1.2.3. Reviewer

B. Smith: IFIPSEC2017, Africacrypt 2017, WCC 2017, Asiacrypt 2017, Eurocrypt 2017, MACIS 2017, PKC 2018

J. Lavauzelle: MACIS 2017

A. Couvreur: Crypto 2017, Eurocrypt 2017, ISIT 2017.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

F. Morain is member of the editorial board of the *Applicable Algebra in Engineering, Communication and Computing*, Springer.

9.1.3.2. Reviewer - Reviewing Activities

B. Smith: Theory of Computing Systems, Springer Women in Mathematics, Research in Number Theory, IEEE Transactions on Information Theory, Journal of Cryptographic Engineering.

A. Couvreur: IEEE Transactions on Information Theory, IEEE Transactions on Communication, Journal of Number Theory, SIAM Journal on Applied Algebra and Geometry.

9.1.4. Invited Talks

B. Smith was an invited speaker at the annual FMF Symposium, a public science event at Universiteit Groningen (Groningen, NL, November 2017)

B. Smith was an invited speaker in the SIAM Applied Algebraic Geometry minisymposium on Applications of Computational Algebraic Geometry to Cryptology (Atlanta, USA, August 2017).

B. Smith was an invited speaker at the FoCM workshop on Computational Number Theory (Barcelona, ES, July 2017)

B. Smith was an invited speaker at the Summer School on Real-World Crypto and Privacy (Sibenik, HR, June 2017)

B. Smith was an invited speaker at JeudiX, a public science outreach event of École polytechnique (Paris, January 2017)

9.1.5. Animation of Seminars

- D. Augot is member of the scientific committee of the CCA seminar, "Codage, Cryptographie et Algorithms", https://cca.inria.fr
- D. Augot, with Bernadette Charron-Bost, is heading the scientific committee of the Blocksem seminar at Polytechnique, on blockchains, http://www.lix.polytechnique.fr/blocksem
- D. Augot, with Fabrice Le Fessant, organised the Open Source Spring on blockchains http://www.open-source-innovation-spring.org/

9.1.6. Research Administration

F. Morain is vice-head of the Département d'informatique of Ecole Polytechnique.

F. Morain is member of the Board of Master Parisien de Recherche en Informatique (MPRI).

A. Couvreur is member of LIX's *Conseil de laboratoire*.

B. Smith was the International Correspondant for CRI Saclay.

B. Smith was a member of the COST-GTRI.

D. Augot is elected member of the "conseil académique consultatif" de Paris-Saclay.

D. Augot was in the "comité de sélection" for a "maître de conférences" position in Grenoble

D. Augot was heading the "comité de sélection" for a "maître de conférences" position in Rouen

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence:

B. Smith, Computer Programming (CSE101), 23h EqTD, L1, École polytechnique, France

J. Lavauzelle, 1I001, *Éléments de programmation*, tutorial class (17.5h equiv TD), L1, Université Pierre et Marie Curie

J. Lavauzelle, 2I003, *Initiation à l'algorithmique*, tutorial class (47.5h equiv TD), L2, Université Pierre et Marie Curie

A. Couvreur and E. Barelli, INF411, "Les bases de la programmation et de l'algorithmique", 21.3h (equiv TD), 2nd year (L3), Ecole Polytechnique, France.

E. Barelli, INF311, "Introduction à l'informatique", 26.7h(equiv TD), 1st year, Ecole Polytechnique, France.

Master:

B. Smith, Advanced Cryptology (INF568), 55h EqTD, M1, École polytechnique, France

B. Smith and F. Morain, Algorithmes Arithmétiques pour la Cryptologie (2-12-2), 20h EqTD, M2, Master Parisien de Recherche en Informatique (MPRI), France

A. Couvreur and F. Morain, Introduction to Cryptology (INF558), 40h, M1, École polytechnique, France

A. Couvreur, Error Correcting Codes and Applications to Cryptography, (2-13-2), 15h, M2, MPRI, FRANCE

Master 2 intern

- D. Augot was the director of Rémi Clarisse internship on the Chor-Rivest cryptosystem

Students project

- D. Augot was managing two groups of polytechniques students on their own project: one about a voting system based on homomorphic encryption (with CEA List), the second about a medical kidney exchange scheme secured and enforced by the Hyperledger/fabric blockchain (with Orange)

9.2.2. Supervision

PhD : Cyril Hugounenq, Volcans et calcul d'isogénies, Université Paris Saclay, 25/09/2017, F. Morain& L. Goubin & L. De Feo.

9.2.3. Juries

- D. Augot
  - examinator of the PhD defense of Sarah Kamel, "Sécurité pour les réseaux sans fil", le 10 mars 2017 (Télécom Paris Tech)
  - examinator of the PhD defense of Francisco Vial-Prado, "Contributions to the design and analysis of fully homomorphic encryption schemes, le 12 juin 2017 (Université Versailles Saint-Quentin)
  - examinator of the PhD defense of Vlad Dragoi "Approche algébrique pour l'étude et la résolution de problèmes algorithmiques issus de la cryptographie et de la théorie des codes", le 6 juillet 2017 (University of Rouen).
  - examinator of the PhD defense of Mohamed A. M. Saeed Taha "Algebraic Approach for Code Equivalence", le 18 décembre 2017 (University of Rouen).
- A. Couvreur
  - PhD : Hervé Talé Kalachi (University of Rouen).
  - Agrégation de Mathématiques.

9.3. Popularization

- A. Couvreur gave the *Conférence inaugurale* of the *Semaine des mathématiques* in the accadémie de Créteil: *Cryptographie, le langage des secrets*.

10. Bibliography

Major publications by the team in recent years

[1] D. AUGOT, M. FINIASZ.*Direct Construction of Recursive MDS Diffusion Layers using Shortened BCH Codes*, in "21st International Workshop on Fast Software Encryption, FSE 2014", London, United Kingdom, C. CID, C. RECHBERGER (editors), springer, March 2014, https://hal.inria.fr/hal-01044597.

[2] A. COUVREUR, I. MÁRQUEZ-CORBELLA, R. PELLIKAAN.*A Polynomial Time Attack against Algebraic Geometry Code Based Public Key Cryptosystems*, in "Information Theory (ISIT), 2014 IEEE International Symposium on", Honolulu, United States, IEEE, June 2014, p. 1446-1450 [*DOI :* 10.1109/ISIT.2014.6875072], https://hal.archives-ouvertes.fr/hal-00937476.

[3] A. COUVREUR, A. OTMANI, J.-P. TILLICH.*Polynomial Time Attack on Wild McEliece Over Quadratic Extensions*, in "EUROCRYPT 2014", Copenhagen, Denmark, May 2014, p. 17-39, https://hal.archives-ouvertes.fr/hal-00931774.

[4] P. LEBACQUE, A. ZYKIN.*On the Number of Rational Points of Jacobians over Finite Fields*, in "Acta Arith.", 2015, vol. 169, p. 373–384, https://hal.archives-ouvertes.fr/hal-01081468.

[5] F. MORAIN.*Implementing the asymptotically fast version of the elliptic curve primality proving algorithm*, in "Math. Comp.", 2007, vol. 76, p. 493–505.

[6] B. SMITH.*Isogenies and the discrete logarithm problem in Jacobians of genus 3 hyperelliptic curves*, in "J. of Cryptology", 2009, vol. 22, n^{o} 4, p. 505-529.

[7] B. SMITH.*Families of fast elliptic curves from Q-curves*, in "Advances in Cryptology - ASIACRYPT 2013", Bangalore, India, K. SAKO, P. SARKAR (editors), Lecture Notes in Computer Science, Springer, December 2013, vol. 8269, p. 61-78 [*DOI :* 10.1007/978-3-642-42033-7_4], https://hal.inria.fr/hal-00825287.

Publications of the year

Articles in International Peer-Reviewed Journal

[8] D. AUGOT, P. LOIDREAU, G. ROBERT.*Generalized Gabidulin codes over fields of any characteristic*, in "Designs, Codes and Cryptography", 2017, https://arxiv.org/abs/1703.09125 [*DOI :* 10.1007/s10623-017-0425-6], https://hal.archives-ouvertes.fr/hal-01503212.

[9] C. COSTELLO, B. SMITH.*Montgomery curves and their arithmetic: The case of large characteristic fields*, in "Journal of Cryptographic Engineering", 2017, https://arxiv.org/abs/1703.01863 [*DOI :* 10.1007/s13389-017-0157-6], https://hal.inria.fr/hal-01483768.

[10] A. COUVREUR, A. OTMANI, J.-P. TILLICH.*Polynomial Time Attack on Wild McEliece Over Quadratic Extensions*, in "IEEE Transactions on Information Theory", January 2017, vol. 63, n^{o} 1, p. 404–427 [*DOI :* 10.1109/TIT.2016.2574841], https://hal.inria.fr/hal-01661935.

[11] C. RITZENTHALER, R. LERCIER, F. ROVETTA, J. SIJSLING, B. SMITH.*Distributions of traces of Frobenius for smooth plane curves over finite fields*, in "Experimental Mathematics", 2017, https://arxiv.org/abs/1510.05601 [*DOI :* 10.1080/10586458.2017.1328321], https://hal.inria.fr/hal-01217995.

International Conferences with Proceedings

[12] D. AUGOT, H. CHABANNE, T. CHENEVIER, W. GEORGE, L. LAMBER.*A User-Centric System for Verified Identities on the Bitcoin Blockchain*, in "International Workshop on Cryptocurrencies and Blockchain Technology - CBT'17", Oslo, Norway, September 2017, https://arxiv.org/abs/1710.02019 , https://hal.inria.fr/hal-01611251.

[13] D. AUGOT, H. CHABANNE, O. CLÉMOT, W. GEORGE.*Transforming face-to-face identity proofing into anonymous digital identity using the Bitcoin blockchain*, in "PST2017 - International Conference on Privacy, Security and Trust", Calgary, Canada, August 2017, 10, https://arxiv.org/abs/1710.02951 , https://hal.inria.fr/hal-01611297.

[14] L. GRÉMY, A. GUILLEVIC, F. MORAIN, E. THOMÉ.*Computing discrete logarithms in $GF(p^6)$*, in "24th Annual Conference on Selected Areas in Cryptography", Ottawa, Canada, August 2017, https://hal.inria.fr/hal-01624662.

[15] J. RENES, B. SMITH.*qDSA: Small and Secure Digital Signatures with Curve-based Diffie–Hellman Key Pairs*, in "ASIACRYPT 2017", Hong Kong, China, IACR, December 2017, https://arxiv.org/abs/1709.03358 , https://hal.inria.fr/hal-01585322.

Conferences without Proceedings

[16] E. BARELLI.*On the security of Some Compact Keys for McEliece Scheme*, in "WCC 2017 - The Tenth International Workshop on Coding and Cryptography", St Petersbourg, Russia, September 2017, p. 1-9, https://hal.inria.fr/hal-01674546.

[17] J. LAVAUZELLE.*Constructions for efficient Private Information Retrieval protocols*, in "WCC 2017 - The Tenth International Workshop on Coding and Cryptography", Saint-Petersbourg, Russia, Inria and SUAI and Skoltech, September 2017, p. 1-12, https://hal.inria.fr/hal-01633469.

Other Publications

[18] C. BACHOC, A. COUVREUR, G. ZÉMOR.*Towards a function field version of Freiman's Theorem*, September 2017, https://arxiv.org/abs/1709.00087 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01584034.

[19] M. BARDET, E. BARELLI, O. BLAZY, R. CANTO TORRES, A. COUVREUR, P. GABORIT, A. OTMANI, N. SENDRIER, J.-P. TILLICH.*BIG QUAKE BInary Goppa QUAsi–cyclic Key Encapsulation*, December 2017, submission to the NIST post quantum cryptography standardization process, https://hal.archives-ouvertes.fr/hal-01671866.

[20] E. BARELLI, P. BEELEN, M. DATTA, V. NEIGER, J. ROSENKILDE.*Two-Point Codes for the Generalized GK Curve*, October 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01535513.

[21] N. COXON.*Fast systematic encoding of multiplicity codes*, April 2017, https://arxiv.org/abs/1704.07083 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01512372.

[22] L. GRÉMY, A. GUILLEVIC, F. MORAIN.*Breaking DLP in $GF(p^5)$ using 3-dimensional sieving*, July 2017, working paper or preprint, https://hal.inria.fr/hal-01568373.

References in notes

[23] A. GUILLEVIC, F. MORAIN, E. THOMÉ.*Solving discrete logarithms on a 170-bit MNT curve by pairing reduction*, in "Selected Areas in Cryptography 2016", St. John's, Canada, R. AVANZI, H. HEYS (editors), Selected Areas in Cryptography 2016, Springer, August 2016, to appear in the Lecture Notes in Computer Science (LNCS), https://hal.inria.fr/hal-01320496.

Project-Team ILDA

Interacting with Large Data

IN COLLABORATION WITH: Laboratoire de recherche en informatique (LRI)

IN PARTNERSHIP WITH:
CNRS

Université Paris-Sud (Paris 11)

RESEARCH CENTER
Saclay - Île-de-France

THEME
Interaction and visualization

Table of contents

Project-Team ILDA

Creation of the Team: 2015 January 01, updated into Project-Team: 2016 December 01

Keywords:

Computer Science and Digital Science:

A3.1.7. - Open data
A3.2.4. - Semantic Web
A3.2.5. - Ontologies
A5.1. - Human-Computer Interaction
A5.2. - Data visualization
A5.5.4. - Animation

Other Research Topics and Application Domains:

B7.1. - Traffic management
B9.4.3. - Physics
B9.4.5. - Data science
B9.5.7. - Geography
B9.7.2. - Open data
B9.9. - Risk management

1. Personnel

Research Scientists

Emmanuel Pietriga [Team leader, Inria, Senior Researcher, HDR]
Caroline Appert [CNRS, Researcher, HDR]
Olivier Chapuis [CNRS, Researcher]

Faculty Member

Anastasia Bezerianos [Univ Paris-Sud, Associate Professor]

External Collaborators

Evanthia Dimara [Inria, PhD Student, co-advised with project-team Aviz]
Bruno Fruchard [Institut Telecom ex GET Groupe des Ecoles des Télécommunications, PhD Student]

Technical Staff

Hande Gözükan [Inria, from Nov 2017]
Dylan Lebout [Inria, from Nov 2017]

PhD Students

Marie Destandau [Inria, from Oct 2017]
Anna Gogolou [Inria]
Maria Jesus Lobo Gunther [Inria, until Dec 2017]
Rafael Morales Gonzalez [Univ Paris-Sud, until Sep 2017]
Arnaud Prouzeau [Univ Paris-Sud, until Dec 2017]
Hugo Romat [Tecknowmetrix]

Post-Doctoral Fellow

Vit Rusnak [Inria]

Administrative Assistant

Alexandra Merlin [Inria]

2. Overall Objectives

2.1. Overall Objectives

In an increasing number of domains, computer users are faced with large datasets, that are often interlinked and organized according to elaborate structures thanks to new data models such as those that are arising with the development of, e.g., the Web of Data. Rather than seeing the inherent complexity of those data models as a hindrance, we aim at leveraging it to design new interactive systems that can better assist users in their data understanding and processing tasks.

These “Data-centric Interactive Systems” aim at providing users with the right information at the right time, presenting it in the most meaningful manner, and letting users efficiently manipulate, edit and share these data with others. This entails minimizing the effort required to retrieve and relate data from relevant sources; displaying data using visual presentation techniques that match the data’s characteristics and the users’ tasks; and providing users with means of interacting with the data that effectively support their train of thought.

Our approach is based on the idea of bringing the fields of Web data management [28] and Human-computer interaction [53], [75] closer together, based on the strong belief that they have the potential to cross-fertilize one another. User interface design is essential to the management and understanding of large, interlinked datasets. Interlinked datasets enriched with even a small amount of semantics have the potential to help create interfaces that let users analyze and manipulate data in a more efficient manner by providing them with, e.g, more relevant query results and giving them efficient means to navigate and relate those results. Our ultimate, long-term goal is to design interactive systems that make it as straightforward to manipulate large webs of data as spreadsheets do for tabular data.

3. Research Program

3.1. Introduction

Our ability to acquire or generate, store, process, interlink and query data has increased spectacularly over the last few years. The corresponding advances are commonly grouped under the umbrella of so called *Big Data*. Even if the latter has become a buzzword, these advances are real, and they are having a profound impact in domains as varied as scientific research, commerce, social media, industrial processes or e-government. Yet, looking ahead, emerging technologies related to what we now call the *Web of Data* (a.k.a the Semantic Web) have the potential to create an even larger revolution in data-driven activities, by making information accessible to machines as semistructured data [27] that eventually becomes actionable knowledge. Indeed, novel Web data models considerably ease the interlinking of semi-structured data originating from multiple independent sources. They make it possible to associate machine-processable semantics with the data. This in turn means that heterogeneous systems can exchange data, infer new data using reasoning engines, and that software agents can cross data sources, resolving ambiguities and conflicts between them [73]. Datasets are becoming very rich and very large. They are gradually being made even larger and more heterogeneous, but also much more useful, by interlinking them, as exemplified by the Linked Data initiative [48].

These advances raise research questions and technological challenges that span numerous fields of computer science research: databases, communication networks, security and trust, data mining, as well as human-computer interaction. Our research is based on the conviction that interactive systems play a central role in many data-driven activity domains. Indeed, no matter how elaborate the data acquisition, processing and storage pipelines are, data eventually get processed or consumed one way or another by users. The latter are faced with large, increasingly interlinked heterogeneous datasets (see, e.g., Figure 1) that are organized according to complex structures, resulting in overwhelming amounts of both raw data and structured information. Users thus require effective tools to make sense of their data and manipulate them.

Figure 1. Linking Open Data cloud diagram from 2007 to 2017 – http://lod-cloud.net

We approach this problem from the perspective of the Human-Computer Interaction (HCI) field of research, whose goal is to study how humans interact with computers and inspire novel hardware and software designs aimed at optimizing properties such as efficiency, ease of use and learnability, in single-user or cooperative work contexts. More formally, HCI is about designing systems that lower the barrier between users' cognitive model of what they want to accomplish, and computers' understanding of this model. HCI is about the design, implementation and evaluation of computing systems that humans interact with [53], [75]. It is a highly multidisciplinary field, with experts from computer science, cognitive psychology, design, engineering, ethnography, human factors and sociology.

In this broad context, ILDA aims at designing interactive systems that display [37], [60], [82] the data and let users interact with them, aiming to help users better *navigate* and *comprehend* large webs of data represented visually, as well as *relate* and *manipulate* them.

Our research agenda consists of the three complementary axes detailed in the following subsections. Designing systems that consider interaction in close conjunction with data semantics is pivotal to all three axes. Those semantics will help drive navigation in, and manipulation of, the data, so as to optimize the communication bandwidth between users and data.

3.2. Semantics-driven Data Manipulation

Participants: Emmanuel Pietriga, Caroline Appert, Marie Destandau, Hugo Romat, Hande Gözükan, Dylan Lebout.

The Web of Data has been maturing for the last fifteen years and is starting to gain adoption across numerous application domains (Figure 1). Now that most foundational building blocks are in place, from knowledge representation, inference mechanisms and query languages [49], all the way up to the expression of data presentation knowledge [68] and to mechanisms like look-up services [81] or spreading activation [44], we need to pay significant attention to how human beings are going to interact with this new Web, if it is to *"reach its full potential"* [45].

Most efforts in terms of user interface design and development for the Web of data have essentially focused on tools for software developers or subject-matter experts who create ontologies and populate them [55], [43]. Tools more oriented towards end-users are starting to appear [34], [36], [50], [51], [54], [63], including the so-called *linked data browsers* [48]. However, those browsers are in most cases based on quite conventional point-and-click hypertext interfaces that present data to users in a very page-centric, web-of-documents manner that is ill-suited to navigating in, and manipulating, webs of data.

To be successful, interaction paradigms that let users navigate and manipulate data on the Web have to be tailored to the radically different way of browsing information enabled by it, where users directly interact with the data rather than with monolithic documents. The general research question addressed in this part of our research program is how to design novel interaction techniques that help users manipulate their data more efficiently. By data manipulation, we mean all low-level tasks related to manually creating new content, modifying and cleaning existing content, merging data from different sources, establishing connections between datasets, categorizing data, and eventually sharing the end results with other users; tasks that are currently considered quite tedious because of the sheer complexity of the concepts, data models and syntax, and the interplay between all of them.

Our approach is based on the conviction that there is a strong potential for cross-fertilization, as mentioned earlier: on the one hand, user interface design is essential to the management and understanding of webs of data; on the other hand, interlinked datasets enriched with even a small amount of semantics can help create more powerful user interfaces, that provide users with the right information at the right time.

We envision systems that focus on the data themselves, exploiting the underlying *semantics and structure* in the background rather than exposing them – which is what current user interfaces for the Web of Data often do. We envision interactive systems in which the semantics and structure are not exposed directly to users, but serve as input to the system to generate interactive representations that convey information relevant to the task at hand and best afford the possible manipulation actions.

Relevant publications by team members this year: [21], [15].

3.3. Generalized Multi-scale Navigation

Participants: Caroline Appert, Anastasia Bezerianos, Olivier Chapuis, Emmanuel Pietriga, Marie Destandau, Anna Gogolou, Maria Jesus Lobo Gunther, Arnaud Prouzeau, Vit Rusnak.

The foundational question addressed here is what to display when, where and how, so as to provide effective support to users in their data understanding and manipulation tasks. ILDA targets contexts in which workers have to interact with complementary views on the same data, or with views on different-but-related datasets, possibly at different levels of abstraction. Being able to combine or switch between representations of the data at different levels of detail and merge data from multiple sources in a single representation is central to many scenarios. This is especially true in both of the application domains we consider: mission-critical systems (e.g., natural disaster crisis management) and the exploratory analysis of scientific data (e.g., correlate theories and heterogeneous observational data for an analysis of a given celestial body in Astrophysics).

A significant part of our research over the last ten years has focused on multi-scale interfaces. We designed and evaluated novel interaction techniques, but also worked actively on the development of open-source UI toolkits for multi-scale interfaces (http://zvtm.sf.net). These interfaces let users navigate large but relatively homogeneous datasets at different levels of detail, on both workstations [71], [30], [67], [66], [65], [31], [70], [29], [72] and wall-sized displays [62], [57], [69], [61], [33], [39], [38]. This part of the ILDA research program is about extending multi-scale navigation in two directions: 1. Enabling the representation of multiple, spatially-registered but widely varying, multi-scale data layers in Geographical Information Systems (GIS); 2. Generalizing the multi-scale navigation paradigm to interconnected, heterogeneous datasets as found on the Web of Data.

The first research problem is mainly investigated in collaboration with IGN in the context of ANR project MapMuxing (Section 9.2.1), which stands for *multi-dimensional map multiplexing*. Project MapMuxing aims at going beyond the traditional pan & zoom and overview+detail interface schemes, and at designing and evaluating novel cartographic visualizations that rely on high-quality generalization, *i.e.*, the simplification of geographic data to make it legible at a given map scale [78], [79], and symbol specification. Beyond project MapMuxing, we are also investigating multi-scale multiplexing techniques for geo-localized data in the specific context of ultra-high-resolution wall-sized displays, where the combination of a very high pixel density and large physical surface (Figure 2) enable us to explore designs that involve collaborative interaction and physical navigation in front of the workspace. This is work done in cooperation with team Massive Data at Inria Chile.

The second research problem is about the extension of multi-scale navigation to interconnected, heterogeneous datasets. Generalization has a rather straightforward definition in the specific domain of geographical information systems, where data items are geographical entities that naturally aggregate as scale increases. But it is unclear how generalization could work for representations of the more heterogeneous webs of data that we consider in the first axis of our research program. Those data form complex networks of resources with multiple and quite varied relationships between them, that cannot rely on a single, unified type of representation (a role played by maps in GIS applications).

Addressing the limits of current generalization processes is a longer-term, more exploratory endeavor. Here again, the machine-processable semantics and structure of the data give us an opportunity to rethink how users navigate interconnected heterogeneous datasets. Using these additional data, we investigate ways to generalize the multi-scale navigation paradigm to datasets whose layout and spatial relationships can be much richer and much more diverse than what can be encoded with static linear hierarchies as typically found today in interfaces for browsing maps or large imagery. Our goal is thus to design and develop highly dynamic and versatile multi-scale information spaces for heterogeneous data whose structure and semantics are not known in advance, but discovered incrementally.

Relevant publications by team members this year: [17].

3.4. Novel Forms of Input for Groups and Individuals

Participants: Caroline Appert, Anastasia Bezerianos, Olivier Chapuis, Emmanuel Pietriga, Rafael Morales Gonzalez, Arnaud Prouzeau, Eleonore Bartenlian, Reyhaneh Raissi, Bruno Fruchard.

Analyzing and manipulating large datasets can involve multiple users working together in a coordinated manner in multi-display environments: workstations, handheld devices, wall-sized displays [33]. Those users work towards a common goal, navigating and manipulating data displayed on various hardware surfaces in a coordinated manner. Group awareness [47], [26] is central in these situations, as users, who may or may not be co-located in the same room, can have an optimal individual behavior only if they have a clear picture of what their collaborators have done and are currently doing in the global context. We work on the design and implementation of interactive systems that improve group awareness in co-located situations [56], making individual users able to figure out what other users are doing without breaking the flow of their own actions.

In addition, users need a rich interaction vocabulary to handle large, structured datasets in a flexible and powerful way, regardless of the context of work. Input devices such as mice and trackpads provide a limited number of input actions, thus requiring users to switch between modes to perform different types of data manipulation and navigation actions. The action semantics of these input devices are also often too much dependent on the display output. For instance, a mouse movement and click can only be interpreted according to the graphical controller (widget) above which it is moved. We focus on designing powerful input techniques based upon technologies such as tactile surfaces (supported by UI toolkits developed in-house), 3D motion tracking systems, or custom-built controllers [59] *to complement (rather than replace) traditional input devices* such as keyboards, that remain the best method so far for text entry, and indirect input devices such as mice or trackpads for pixel-precise pointing actions.

The input vocabularies we investigate enable users to navigate and manipulate large and structured datasets in environments that involve multiple users and displays that vary in their size, position and orientation [33], [46], each having their own characteristics and affordances: wall displays [62], [84], workstations, tabletops [64], [42], tablets [7], [80], smartphones [83], [40], [76], [77], and combinations thereof [2], [10], [61], [33].

We aim at designing rich interaction vocabularies that go far beyond what current touch interfaces offer, which rarely exceeds five gestures such as simple slides and pinches. Designing larger gesture vocabularies requires identifying discriminating dimensions (e.g., the presence or absence of anchor points and the distinction between internal and external frames of reference [7]) in order to structure a space of gestures that interface designers can use as a dictionary for choosing a coherent set of controls. These dimensions should be few and simple, so as to provide users with gestures that are easy to memorize and execute. Beyond gesture complexity, the scalability of vocabularies also depends on our ability to design robust gesture recognizers that will allow users to fluidly chain simple gestures that make it possible to interlace navigation and manipulation actions.

We also study how to further extend input vocabularies by combining touch [7], [83], [64] and mid-air gestures [62] with physical objects [52], [74], [59] and classical input devices such as keyboards to enable users to input commands to the system or to involve other users in their workflow (request for help, delegation, communication of personal findings, etc.) [35], [58]. Gestures and objects encode a lot of information in their shape, dynamics and direction, that can be directly interpreted in relation with the user, independently from the display output. Physical objects can also greatly improve coordination among actors for, e.g., handling priorities or assigning specific roles.

Relevant publications by team members this year: [11], [23], [22], [20], [24], [15], [14], [18].

4. Application Domains

4.1. Mission-critical systems

Mission-critical contexts of use include emergency response & management, and critical infrastructure operations, such as public transportation systems, communications and power distribution networks, or the operations of large scientific instruments such as particle accelerators and astronomical observatories. Central to these contexts of work is the notion of situation awareness [26], i.e., how workers perceive and understand elements of the environment with respect to time and space, such as maps and geolocated data feeds from the field, and how they form mental models that help them predict future states of those elements. One of the main challenges is how to best assist subject-matter experts in constructing correct mental models and making informed decisions, often under time pressure. This can be achieved by providing them with, or helping them efficiently identify and correlate, relevant and timely information extracted from large amounts of raw data, taking into account the often cooperative nature of their work and the need for task coordination. With this application area, our goal is to investigate novel ways of interacting with computing systems that improve collaborative data analysis capabilities and decision support assistance in a mission-critical, often time-constrained, work context.

Relevant publications by team members this year: [24], [15], [18], [17], [25].

4.2. Exploratory analysis of scientific data

Many scientific disciplines are increasingly data-driven, including astronomy, molecular biology, particle physics, or neuroanatomy. While making the right decision under time pressure is often less of critical issue when analyzing scientific data, at least not on the same temporal scale as truly time-critical systems, scientists are still faced with large-to-huge amounts of data. No matter their origin (experiments, remote observations, large-scale simulations), these data are difficult to understand and analyze in depth because of their sheer size and complexity. Challenges include how to help scientists freely-yet-efficiently explore their data, keep a trace of the multiple data processing paths they considered to verify their hypotheses and make it easy to backtrack, and how to relate observations made on different parts of the data and insights gained at different moments during the exploration process. With this application area, our goal is to investigate how data-centric interactive systems can improve collaborative scientific data exploration, where users' goals are more open-ended, and where roles, collaboration and coordination patterns [47] differ from those observed in mission-critical contexts of work.

Relevant publications by team members last year: [8].

5. Highlights of the Year

5.1. Highlights of the Year

- Caroline Appert was papers co-chairs for the 2017 ACM CHI Conference on Human Factors in Computing Systems, the flagship conference in HCI, with more than 2,400 submissions in 2017.

- Caroline Appert joined the editorial board of ACM ToCHI (Transactions on Computer-Human Interaction), one of the two top journals in HCI.

6. New Software and Platforms

6.1. Smarties

FUNCTIONAL DESCRIPTION: The Smarties system provides an easy way to add mobile interactive support to collaborative applications for wall displays.

It consists of (i) a mobile interface that runs on mobile devices for input, (ii) a communication protocol between the mobiles and the wall application, and (iii) libraries that implement the protocol and handle synchronization, locking and input conflicts. The library presents the input as an event loop with callback functions and handles all communication between mobiles and wall application. Developpers can customize the mobile interface from the wall application without modifying the mobile interface code.

On each mobile we find a set of cursor controllers associated with keyboards, widgets and clipboards. These controllers (pucks) can be shared by multiple collaborating users. They can control simple cursors on the wall application, or specific content (objects or groups of them). The developper can decide the types of widgets associated to pucks from the wall application side.

- Contact: Olivier Chapuis
- URL: http://smarties.lri.fr/

6.2. ZVTM

Zoomable Visual Transformation Machine
KEYWORDS: Big data - Visualization - Data visualization - Information visualization - Graph visualization
FUNCTIONAL DESCRIPTION: ZVTM is a toolkit enabling the implementation of multi-scale interfaces for interactively navigating in large datasets displayed as 2D graphics.

ZVTM is used for browsing large databases in multiple domains: geographical information systems, control rooms of complex facilitites, astronomy, power distribution systems.

The toolkit also enables the development of applications running on ultra-high-resolution wall-sized displays.

- Participants: Arnaud Prouzeau, Can Liu, Caroline Appert, Hande Gozukan, Maria Jesus Lobo Gunther and Olivier Chapuis
- Contact: Emmanuel Pietriga
- URL: http://zvtm.sf.net

6.3. Platforms

6.3.1. Platform: WILDER

Ultra-high-resolution wall-sized displays [33] feature a very high pixel density over a large physical surface. Such platforms have properties that make them well-suited to the visualization of very large datasets. They can represent the data with a high level of detail while at the same time retaining context: users can transition from an overview of the data to a detailed view simply by physically moving in front of the wall display. Wall displays also offer good support for collaborative work, enabling multiple users to simultaneously visualize and interact with the displayed data. To make them interactive, wall-sized displays are increasingly coupled with input devices such as touch frames, motion-tracking systems and wireless multitouch devices, in order to enable multi-device and multi-user interaction with the displayed data. Application areas for such visualization platforms range from the monitoring of complex infrastructures and crisis management situations to tools for the exploratory visualization of scientific data.

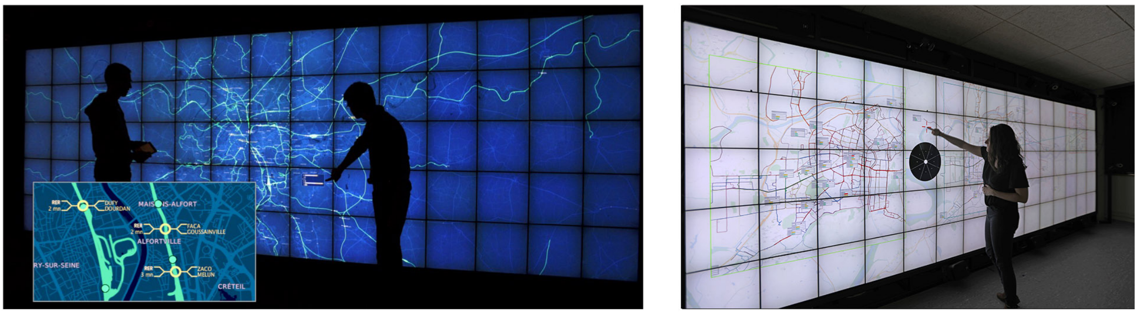


Figure 2. Geovisualization applications running on the WILDER platform. Real-time monitoring of railroad traffic in France (left), real-time monitoring of mobile sensors measuring air quality in Korean cities (right).

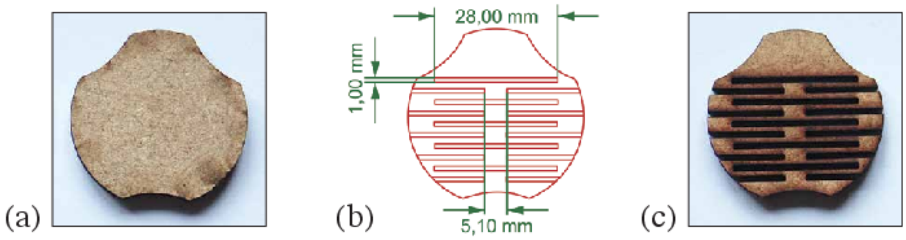
WILDER is the latest ultra-high-resolution wall-sized display set up at Inria Saclay, and is one of the nodes of the Digiscope EquipEx. We use this platform for multiple projects, both fundamental HCI research, and research and development activities for specific application areas such as geographical informations systems (Figure 2) and astronomy.

WILDER was used in the projects that led to the following publications this year: [22], [24], [18].

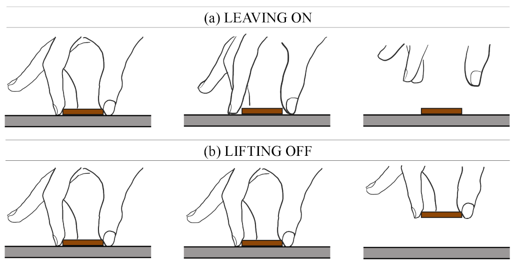
7. New Results

7.1. Gestures and Tangibles

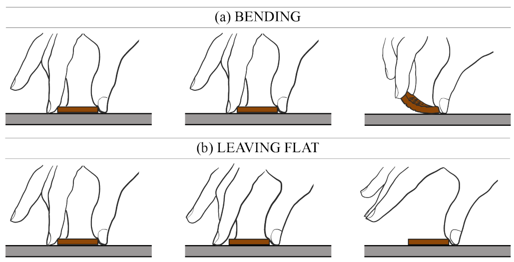
- As a follow-up to our work on TouchTokens [6], we investigated a way to augment the expressiveness of passive tokens for tangible interaction. This work was published at CHI 2017 [23]. TouchTokens are passive tokens that can be recognized on any capacitive surface based on the spatial configuration of the fingers that hold them. However, interaction with these tokens is confined to the basic two-state model of touch interaction as the system only knows the tokens' position and cannot detect tokens that are not touched. We increased the expressive power of TouchTokens by introducing laser-cut lattice hinges in their design, so as to make them flexible (Figure 3). A new recognizer, that analyzes the micro-movements of the fingers that hold the tokens, enables the system to detect when a token is left on the surface rather than taken off it. It can also detect bend events that can be mapped to command triggers, and a squeezed state that can be used for quasi-modal interaction.
- With MarkPad, presented at CHI 2017 [20] and demoed at IHM 2017, we propose a novel interaction technique taking advantage of the touchpad. MarkPad allows creating a large number of size-dependent gestural shortcuts that can be spatially organized as desired by the user. It relies on the idea of using visual or tactile marks on the touchpad or a combination of them. Gestures start from a mark on the border and end on another mark anywhere (see Figure 4). MarkPad does not conflict with standard interactions and provides a novice mode that acts as a rehearsal of the expert mode. A study showed that an accuracy of 95% could be achieved for a dense configuration of tactile and/or visual marks allowing 680 possible gestures, more than all existing techniques with a comparable input channel. Performance was 5% lower in a second study where the marks were only on the borders, and subjective results suggest that a mixed interface (borders with tactile marks and center with visual marks) is a promising solution. A working prototype is freely available at http://brunofruchard.com/markpad.html.



Making a TouchToken flexible: (a) original, rigid TouchToken (circle, 4cm in diameter), (b) schematics of lattice-hinges, (c) flexible TouchToken.



Micro-movements when leaving a token *on* the surface (a), and when lifting it *off* (b)



Micro-movements when (a) bending a token, and when leaving it flat (b)

Figure 3. Flexible TouchTokens

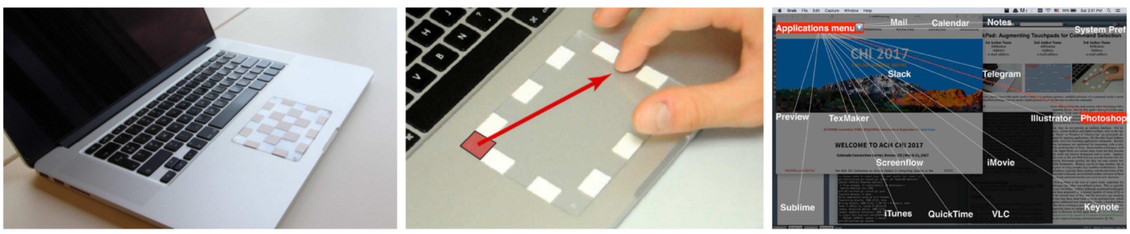


Figure 4. Tactile or visual marks on the touchpad help performing gestures: (Left) Dense configuration of tactile marks, (Middle) Light configuration with marks only on the borders, (Right) Example of a menu in novice mode. This menu and the selected shortcut on its right side correspond to the red area and the red gesture line in the middle picture.

7.2. Interacting with Linked Data and the Semantic Web

- As part of the team’s novel research theme on Semantics-Driven Data Manipulation 3.2, Emmanuel Pietriga worked jointly with colleagues from Linköping University on a visualization technique for the comparative evaluation of ontology alignments produced by different algorithms, that was published at the International Semantic Web Conference [21]. Ontology alignment is an area of active research where many algorithms and approaches are being developed. Their performance is usually evaluated by comparing the produced alignments to a reference alignment in terms of precision, recall and F-measure. These measures, however, only provide an overall assessment of the quality of the alignments, but do not reveal differences and commonalities between alignments at a finer-grained level such as, e.g., regions or individual mappings. Furthermore, reference alignments are often unavailable, which makes the comparative exploration of alignments at different levels of granularity even more important. Making such comparisons efficient calls for a human-in-the-loop approach, best supported through interactive visual representations of alignments. Our approach extended previous work by Inria on Matrix Cubes [32], used for visualizing dense dynamic networks. We identified use cases for ontology alignment evaluation that could benefit from interactive visualization, and then detailed how *Alignment Cubes* could support interactive exploration of multiple ontology alignments. We then showed how alignment cubes could support common tasks identified in these use cases.

7.3. Wall Displays

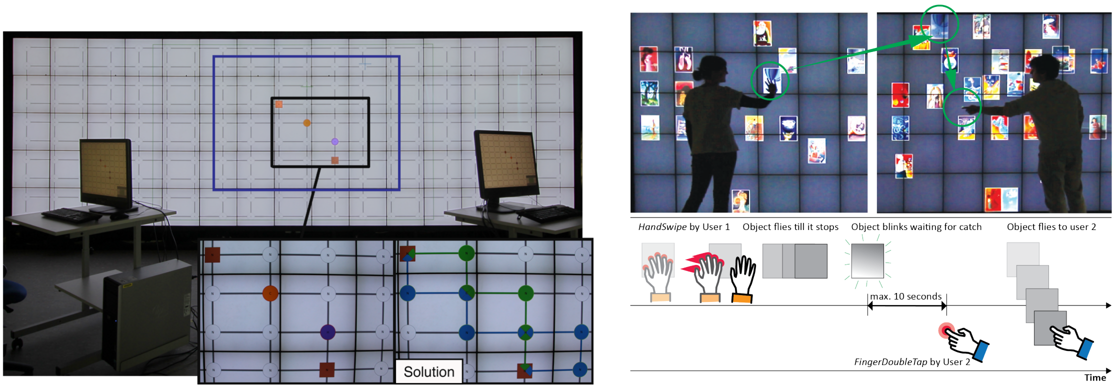


Figure 5. (Left) The setup for comparing a wall display with two desktops. Participants had to construct paths between the two brown nodes, that only crossed at the other colored nodes. The lower part of the image shows the closeup of a possible solution. (Right) An example of cooperative a gesture: throw-and-catch.

Ultra-high-resolution wall displays feature a very high pixel density over a large physical surface, which makes them well-suited to the collaborative, exploratory visualization of large datasets (see Section 6.3.1). We have continued working on the design, implementation and evaluation of interactive visualization techniques for such ultra-high-resolution wall-sized displays, focusing, in some of these projects, on the collaboration between users who perform different data manipulation and analysis tasks.

- We first studied if wall displays indeed provide advantages over more classic collaboration setups, such as multiple desktops [24]. Very few studies that empirically assess the differences of collaboration in front of a shared display compared to a non-shared setup, such as multiple desktops with a common view. We compared the use of the wall compared to two desktops, when pairs of users learn to perform a path-planning task (see Figure 5-(Left)). Path planning tasks are common in critical situations (e.g., rerouting resources). We focused on learning, to approach exceptional and unexpected events in critical systems. Our results did not indicate a significant difference in learning time between the two setups, but found that participants adopted different task strategies and that quality was more consistent in the wall setup.

- We also continued our work on shared interaction techniques (see [56]). Multi-touch wall-sized displays afford collaborative exploration of large datasets and re-organization of digital content. However, standard touch interactions, such as dragging to move content, do not scale well to large surfaces and were not designed to support collaboration, such as passing an object around. With *CoReach* [22], published at CHI 2017, we introduce a set of collaborative gestures that combine input from multiple users in order to manipulate content, facilitate data exchange and support communication (see Figure 5-(Right) for an example). We conducted an observational study to inform the design of *CoReach*, and a controlled study showing that it reduced physical fatigue and facilitated collaboration when compared with traditional multi-touch gestures. A final study assessed the value of also allowing input through a handheld tablet to manipulate content from a distance.
- We also studied more explicitly how interaction techniques affect collaboration. We investigated how pairs explore graphs on a touch enabled wall-display [18], using two selection techniques adapted for collaboration: a basic localized selection, and a propagation selection technique that uses the idea of diffusion/transmission from an origin node. Results from this work were accepted for publication in 2016 (and was part of the previews report), but the work appeared in print this year.

7.4. Visualization

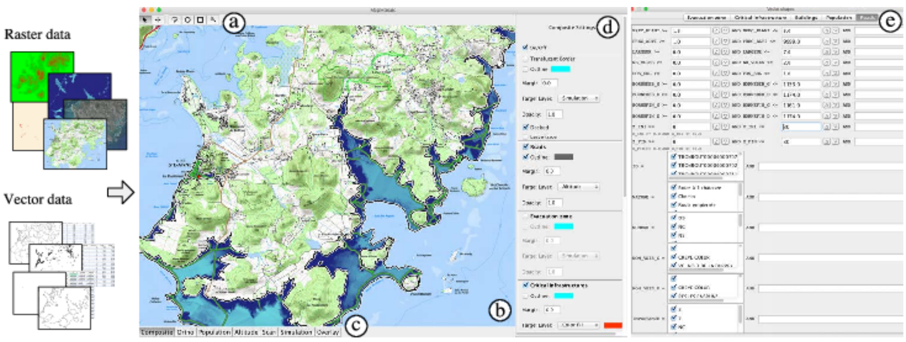


Figure 6. A proof-of-concept implementation of the MapMosaic dynamic compositing model and interaction techniques. (a) Toolbar to navigate the map, and to create & select areas. (b) Map viewer. (c) Access to individual layers. (d) Compositing area inspector. (e) Query builder for compositing region filters.

- In the context of ANR-funded collaborative project MapMuxing (see Section 9.2.1), we investigated novel dynamic map compositing techniques for geovisualization. GIS software applications and other mapping tools enable users to correlate data from multiple layers and gain insight from the resulting visualizations. However, most of these applications only feature basic, monolithic layer compositing techniques. These techniques do not always support users effectively in their tasks, as we observed during interviews with GIS experts at IGN (French national cartographic institute). We designed MapMosaic (Figure 6), a novel approach based on dynamic visual compositing that enables users to interactively create and manipulate local composites of multiple vector and raster map layers, taking into account the semantics and attribute values of objects and fields in the compositing process. We evaluated MapMosaic's interaction model against that of QGIS (a widely used desktop GIS) and MAPublisher (a professional cartography tool) using the Cognitive Dimensions framework, showing that MapMosaic's model is more flexible and can support users more effectively in their tasks. Feedback obtained from experts further confirmed the potential of this highly dynamic approach to map layer compositing.

- We also explored how different interactive visualizations of multidimensional datasets can affect how we make decisions [15]. We evaluated three elementary visualizations: parallel coordinates, scatterplot matrices and tabular visualizations. Our method consisted in first giving participants low-level analytic tasks, in order to ensure that they properly understood the visualizations and their interactions. Participants were then given multi-attribute choice tasks consisting of choosing holiday packages. We assessed decision support through multiple objective and subjective metrics, including a decision accuracy metric based on the consistency between the choice made and self-reported preferences for attributes. We found the three visualizations to be comparable on most metrics, with a slight advantage for tabular visualizations. In particular, tabular visualizations allow participants to reach decisions faster. Our results also suggest that indirect methods for assessing choice confidence may allow to better distinguish between visualizations than direct ones. Related to this topic, is our previous work on studying how biases can affect our decision making when using visualizations [14], work that was accepted in 2016 (and thus was part of last year's report) but appeared in print this year.
- Beyond the actual interactive visualizations themselves, we studied how framing the questions to participants may affect the results of evaluating visualizations [19]. More specifically we explored the effects of providing task context when evaluating visualization tools in crowdsourced studies. We gave participants abstract information visualization tasks without any context; tasks where we added semantics to the dataset; and tasks with two types of backstory narratives: an analytic narrative and a decision-making narrative. We did not find evidence that adding data semantics increases accuracy, but that it increases attention and provides subjective benefits in terms of confidence, perceived easiness, task enjoyability and perceived usefulness of the visualization. Interestingly, we also found that backstory narratives, often used to motivate study participants, can even decrease accuracy.
- Finally, we are interested in understanding how people understand more general multidimensional visualisations. We mention here again work with colleagues both from University of Konstanz [16] on a review of multidimensional visualizations in the forms of glyphs; and with colleagues from INRA on a mixed initiative system that aids navigation of complex multi-dimensional datasets [41]. Both these results were accepted for publication in 2016 (and were part of the previous report), but the work appeared in print this year.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- Tecknowmetrix (TKM): ANRT/CIFRE PhD (Hugo Romat), 3 years, starting June 2016.

9. Partnerships and Cooperations

9.1. Regional Initiatives

ScaleGest. Surface Gestures for Advanced Graphical Interfaces: Which Gesture for What. (2014-2017) Funded by Digiteo. In collaboration with Telecom ParisTech: **109Keuros**. Participants: Caroline Appert (PI), Rafael Morales Gonzalez, Emmanuel Pietriga.

The project aims at designing gesture-based interaction for expert users who navigate and manipulate large datasets. In the context of advanced graphical applications, the number of gestures should be large-enough to cover the set of controls (*i.e.*, commands and parameter settings) but remain simple-enough to avoid exceeding human abilities. Making gesture-based interaction scale with graphical applications' growing complexity can be achieved only by understanding the foundational aspects of this input modality. This project is about characterizing and structuring both the space of application controls and the space of surface gestures in order to establish guidelines for appropriate control-gesture mappings. It is also about the definition of a sound and systematic evaluation methodology that will serve as a reference benchmark for evaluating these mappings. The resulting control-gesture mappings are demonstrated in the specific application domains of cartography and astronomy.

9.2. National Initiatives

9.2.1. ANR

MapMuxing - Multi-dimensional Map Multiplexing. (2014-2018) Funded by the French National Research Agency (ANR). In collaboration with IGN (Institut National de l'Information Géographique et Forestière): **208Keuros**/499Keuros. Participants: Emmanuel Pietriga (PI), Caroline Appert, Olivier Chapuis, Maria Jesus Lobo Gunther, Vit Rusnak. http://mapmuxing.ign.fr

The project explores novel ways of combining different maps and data layers into a single cartographic representation, and investigates novel interaction techniques for navigating in it. The project aims at going beyond the traditional pan & zoom and overview+detail interface schemes, and at designing and evaluating novel cartographic visualizations that rely on high-quality generalization, *i.e.*, the simplification of geographic data to make it legible at a given map scale, and symbol specification.

9.2.2. Inria - Ministère de la Culture

Visual Exploration of Linked Data on BnF's data portal (2017-2018) Funded by the French Ministère de la Culture and Inria. **65Keuros**. Participants: Emmanuel Pietriga (PI), Caroline Appert, Hande Gözükan, Marie Destandau, Dylan Lebout.

The project explores novel ways of visually navigating the data exposed by the Bibliothèque Nationale de France as linked data on http://data.bnf.fr.

9.2.3. Inria Project Lab

ILDA participates to Inria Project Lab iCODA : Data Journalism : knowledge-mediated Content and Data Interactive Analytics, that started in 2017. A key issue in data science is the design of algorithms that enable analysts to infer information and knowledge by exploring heterogeneous information sources, structured data, or unstructured content. With journalism data as a landmark use-case, iCODA aims to develop the scientific and technological foundation for collaborative, heterogeneous data analysis, guided by formalized, user-centric knowledge. The project relies on realistic scenarios in data-journalism to assess the contribution of the project to this area. iCODA is at the crossroads of several research areas (content analysis, data management, knowledge representation, visualization) and is part of a club of partners of the world of the press. Equipes-projets Inria : Graphik, Ilda, Linkmedia, Cedar. Press partners: Le Monde, OuestFrance, AFP. Participants: Anastasia Bezerianos (PI) and Emmanuel Pietriga.

9.3. European Initiatives

9.3.1. Collaborations with Major European Organizations

- Deutsches Elektronen-Synchrotron (DESY): Scientific collaboration on the design and implementation of user interfaces for array operations monitoring and control for the Cherenkov Telescope Array (CTA) project, to be built in the Canary Islands (Spain) and in the Atacama desert (Chile).

9.4. International Initiatives

9.4.1. Inria International Labs

Inria Chile / CIRIC. From 2012 to 2015, Emmanuel Pietriga was the scientific leader of the Massive Data team at Inria Chile, working on projects in collaboration with the ALMA radio-telescope and the Millenium Institute of Astrophysics. He is now scientific advisor to Inria Chile's visualization lab, and is actively involved in the collaboration between Inria Chile and the LSST on the design and development of user interfaces for operations monitoring and control.

9.4.2. Inria International Partners

9.4.2.1. Informal International Partners

- KISTI (Korea). 2017. We investigated the potential of ultra-high-resolution wall-sized displays for the visualization of stream IOT data in the field of air quality monitoring in large and dense urban areas in Korea. The goal of the project was to design and implement an interactive multi-scale visualization of streamed data collected from vehicles (taxis) equipped with a battery of sensors and geolocation devices. The project focused on how to design effective visualizations that take advantage of the specific characteristics of large surfaces featuring a very high pixel density ; and on how to handle streams of IOT data, in this case the sensor data from all taxis, both live data streams and historical data retrieved from a database.
- University of Konstanz: Daniel Keim and Johannes Fuchs on mapping out the design space for visualization glyphs [16]. Participants: Anastasia Bezerianos.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

- Iftach Sadeh, DESY/CTA Observatory, February 2017

9.5.1.1. Internships

- María Grazia Prato, Inria Chile, October 2017
- Amanda Ibsen, Sebastian Pereira, María Grazia Prato, Inria Chile, June 2017

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- EICS 2018, 10th ACM SIGCHI Symposium on Engineering Interactive Computing Systems: Emmanuel Pietriga (general co-chair)
- CHI 2017, 35th ACM SIGCHI Conference on Human Factors in Computing Systems: Caroline Appert (papers co-chair)

10.1.1.2. Member of the Organizing Committees

- VIS 2017, the IEEE Visualization Conference (SciVis, InfoVis, VAST): Anastasia Bezerianos (Communities co-chair)

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

- CHI 2018, 36th ACM SIGCHI Conference on Human Factors in Computing Systems: Emmanuel Pietriga (SC - subcommittee chair)

10.1.2.2. Member of the Conference Program Committees

- CHI 2018, 36th ACM SIGCHI Conference on Human Factors in Computing Systems: Anastasia Bezerianos
- UIST 2017, 30th ACM User Interface Software and Technologies Symposium: Anastasia Bezerianos
- InfoVis 2017, the IEEE Visualization Conference (SciVis, InfoVis, VAST): Anastasia Bezerianos
- WWW 2018, 27th Web Conference, research track Web Content Analysis, Semantics, and Knowledge: Emmanuel Pietriga
- ISWC 2017, 16th International Semantic Web Conference: Emmanuel Pietriga
- EICS 2017, 9th ACM SIGCHI Symposium on Engineering Interactive Computing Systems: Emmanuel Pietriga
- VOILA @ ISWC 2017, Visualizations and User Interfaces for Ontologies and Linked Data, workshop co-located with ISWC 2017: Emmanuel Pietriga
- Immersive Analytics @ InfoVis 2017: Exploring Future Interaction and Visualization Technologies for Data Analytics, workshop co-located with IEEE InfoVis 2017: Emmanuel Pietriga, Anastasia Bezerianos

10.1.2.3. Reviewer

- ACM CHI 2018, Conference on Human Factors in Computing Systems: Caroline Appert, Olivier Chapuis, Bruno Fruchard, Maria Jesus Lobo Gunther, Arnaud Prouzeau
- ACM UIST 2017, Symposium on User Interface Software and Technology: Caroline Appert, Olivier Chapuis
- ACM ISS 2017, International Conference on Interactive Surfaces and Spaces: Olivier Chapuis
- ACM TEI 2017, International Conference on Tangible, Embedded and Embodied Interaction: Caroline Appert, Bruno Fruchard (WiP), Vit Rusnak (WiP)
- ACM MobileHCI 2017 , International Conference on Human-Computer Interaction with Mobile Devices and Services: Caroline Appert
- GI 2017, Graphics Interface: Olivier Chapuis, Arnaud Prouzeau
- Graph Drawing 2017: Emmanuel Pietriga
- EuroVis EG/VGTC 2017, Conference on Data Visualization: Anastasia Bezerianos
- OzCHI 2017: Vit Rusnak
- IHM 2017, Conference of the Association Francophone d'Interaction Homme-Machine: Olivier Chapuis

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- ACM ToCHI, Transactions on Computer-Human Interaction: Caroline Appert

10.1.3.2. Reviewer - Reviewing Activities

- ACM ToCHI, Transactions on Computer-Human Interaction: Olivier Chapuis
- IEEE TVCG, Transactions on Visualization and Computer Graphics: Olivier Chapuis, Emmanuel Pietriga, Caroline Appert, Anastasia Bezerianos
- IEEE Pervasive Computing: Emmanuel Pietriga
- Future Generation Computer Systems: Vit Rusnak

10.1.4. Invited Talks

- Emmanuel Pietriga and Olivier Chapuis: Monitoring Air Quality in Korea's Metropolises on Ultra-High Resolution Wall-Sized Displays, Asia Data Week 2017, Jeju, South Korea, November 2017

- Caroline Appert: Reconnaître un pattern de points de contact sur une surface tactile – *Multi-Touch* et *Machine Learning*, Journée IHM-IA, Paris (UPMC), France, March 2017.
- Anastasia Bezerianos: Concepts of Information Visualization research, and ties to Data Management, DBTrento, University of Trento, Italy, November 2017.

10.1.5. Leadership within the Scientific Community

- ANR, CES Vice chair - Interaction, Robotique: Caroline Appert

10.1.6. Scientific Expertise

- H2020, Leadership in Enabling and Industrial Technologies - Space RIA: Emmanuel Pietriga
- ANR, CES Vice chair - Interaction, Robotique: Caroline Appert

10.1.7. Research Administration

- Deputy Director of the Laboratoire de Recherche en Informatique (LRI): Olivier Chapuis
- President of Inria Saclay - Île de France's Commission for Technological Development (CDT): Emmanuel Pietriga

10.1.8. Learned societies

- Association Francophone d'Interaction Homme-Machine (AFIHM), in charge of the relation with the SIF: Olivier Chapuis.
- SigCHI Paris Local Chapter, chair: Anastasia Bezerianos.
- SigCHI Paris Local Chapter, vice chair: Caroline Appert.

10.1.9. Hiring committees

- Univ. Paris-Sud hiring committee, Commission Consultative des Spécialistes de l'Université 27ème section (computer science), members: Caroline Appert.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master: Anastasia Bezerianos, Head of EIT masters M1 & M2 HCID, Univ. Paris-Sud

Master: Anastasia Bezerianos, Head of M2 Interaction, Univ. Paris-Sud, Université Paris-Saclay

Master: Anastasia Bezerianos, Programming of Interactive Systems – 21h CM, M1/M2 HCID + M2R Interaction, Univ. Paris-Sud

Master: Anastasia Bezerianos, Mixed Reality and Tangible Interaction – 11h CM, M1/M2 HCID + M2R Interaction, Univ. Paris-Sud

Master: Anastasia Bezerianos, Career Seminar, – 12h CM, M2 HCID + M2R Interaction, Univ. Paris-Sud

Master: Anastasia Bezerianos, Design Project in HCI, – 21h CM, M1 HCID, Univ. Paris-Sud

Master: Anastasia Bezerianos, HCI Project, – 21h CM, M2 HCID + M2 Interaction, Univ. Paris-Sud

Master: Anastasia Bezerianos, Introduction to Java – 18h CM, M1/M2 HCID + M2R Interaction, Univ. Paris-Sud

Master: Emmanuel Pietriga, Data Visualization, 24h CM, M2 Informatique Décisionelle, Univ. Paris-Dauphine.

Master: Caroline Appert, Evaluation of Interactive Systems, 21h CM, M2 Interaction Univ. Paris-Sud / M1-M2 HCID EIT Digital.

Master: Caroline Appert, Advanced Evaluation of Interactive Systems, 21h CM, M2 Interaction Univ. Paris-Sud / M1-M2 HCID EIT Digital.

Master: Arnaud Prouzeau, Programmation des interfaces interactives avancées, 27h, M1 Info Univ. Paris-Sud

Licence: Maria Jesus Lobo Gunther, Algorithmique Informatique Graphique, 27h, L3 Univ. Paris-Sud

Licence: Maria Jesus Lobo Gunther, Programmation des interfaces interactives avancées, 27h, L3 Univ. Paris-Sud

Licence: Maria Jesus Lobo Gunther, Interaction Homme-Machine, 12h TP, L3 Polytech Paris Sud

Licence: Arnaud Prouzeau, Interaction Homme-Machine, 12h TP, L3 Polytech Paris Sud

Licence: Arnaud Prouzeau, Programmation Impérative Avancée en C++, 16.5h TP, L1 MPI, Université Paris-Sud

Licence: Bruno Fruchard, Visualisation, 10h, Télécom Paristech.

Licence: Bruno Fruchard, Interaction Homme-Machine, 16h, Télécom Paristech.

Licence: Bruno Fruchard, Paradigmes de programmation, 12h, Télécom Paristech.

Licence: Bruno Fruchard, Infrastructures et plateformes pour l'informatique répartie, 34h, Télécom Paristech.

10.2.2. Supervision

PhD : Maria Jesus Lobo Gunther, Interaction Techniques for Map Multiplexing, defended on December 5th, 2017. Advisors: Caroline Appert, Emmanuel Pietriga

PhD : Rafael Morales Gonzalez, Surface Gestures for Advanced Graphical Interfaces: Which Gesture for What, defended on October 4th, 2017. Advisors: Caroline Appert, Gilles Bailly, Emmanuel Pietriga

PhD : Arnaud Prouzeau, Collaboration around Wall-Displays in Time Critical and Command and Control Contexts, defended on December 15th, 2017. Advisors: Anastasia Bezerianos, Olivier Chapuis

PhD : Evanthia Dimara, Merging Interactive Visualization and Automated Analysis for Group Decision-Making Involving Large Datasets, defended on November 30th, 2017, Advisors: Pierre Dragicevic, Anastasia Bezerianos

PhD in progress : Marie Destandau, Interactive Visual Exploration of Webs of Data, since October 2017, Advisors: Caroline Appert, Emmanuel Pietriga

PhD in progress : Anna Gogolou, Iterative and expressive querying for big data series, October 2016, Advisors: Anastasia Bezerianos, Themis Palpanas

PhD in progress : Hugo Romat, Visual exploration and interactive manipulation techniques for collections of heterogeneous data and documents, since June 2016, Advisors: Caroline Appert, Emmanuel Pietriga

PhD in progress : Bruno Fruchard, Techniques d'interaction exploitant la mémoire spatiale pour faciliter l'accès rapide aux commandes et aux données, since October 2015, Advisors: Eric Lecolinet, Olivier Chapuis

Master (M2): Dylan Lebout, Visualization of geolocated and temporal data on the Web , May - September 2017, Advisors: Caroline Appert, Emmanuel Pietriga

Master (M2): Reyhaneh Raissi, Multi-user Multi-selection on Tabletop, Mars - September 2017, Advisors: Anastasia Bezerianos, Olivier Chapuis

Master (M1): Eleonore Bartenlian, Tools for customizing tangible interfaces, May - July 2017, Advisors: Caroline Appert

10.2.3. Juries

HDR: Guillaume Touya (IGN/Université Paris-Est): Emmanuel Pietriga (examinateur)

PhD: Justin Mathew (Inria/Université Paris-Saclay): Emmanuel Pietriga (président du jury)

PhD: Lonni Besançon (Inria/Université Paris-Saclay): Emmanuel Pietriga (président du jury)

PhD: Hanaë Rateau (Inria/Université Lille 1): Caroline Appert (examinateur)

PhD: Maxime Guillon (Université Grenoble Alpes): Caroline Appert (examinateur)

PhD: Antoine Lhuillier (ENAC/Université Toulouse III): Caroline Appert (examinateur)

10.3. Popularization

- *Fête de la science*, Université Paris-Sud: Caroline Appert.
- *Science en Direct*, Cité de la Science: Maria Jesus Lobo Gunther.
- *Qu'est-ce qu'un chercheur en IHM ?*, to high-school students visiting Telecom ParisTech: Bruno Fruchard.

11. Bibliography

Major publications by the team in recent years

[1] C. Appert, O. Chapuis, E. Pietriga, M.-J. Lobo.*Reciprocal Drag-and-Drop*, in "ACM Transactions on Computer-Human Interaction", September 2015, vol. 22, n^{o} 6, p. 29:1–29:36 [*DOI :* 10.1145/2785670], https://hal.archives-ouvertes.fr/hal-01185805.

[2] O. Chapuis, A. Bezerianos, S. Frantzeskakis.*Smarties: An Input System for Wall Display Development*, in "CHI '14", Toronto, Canada, ACM, April 2014, p. 2763-2772 [*DOI :* 10.1145/2556288.2556956], https://hal.archives-ouvertes.fr/hal-00979034.

[3] C. Liu, O. Chapuis, M. Beaudouin-Lafon, E. Lecolinet.*CoReach: Cooperative Gestures for Data Manipulation on Wall-sized Displays*, in "Proceedings of the 35th international conference on Human factors in computing systems", Denver, United States, ACM (editor), CHI '17, May 2017 [*DOI :* 10.1145/3025453.3025594], https://hal.archives-ouvertes.fr/hal-01437091.

[4] M.-J. Lobo, C. Appert, E. Pietriga.*MapMosaic: Dynamic Layer Compositing for Interactive Geovisualization*, in "International Journal of Geographical Information Science", May 2017, vol. 31, n^{o} 9, p. 1818 - 1845 [*DOI :* 10.1109/TVCG.2007.70570], https://hal.inria.fr/hal-01562084.

[5] M.-J. Lobo, E. Pietriga, C. Appert.*An Evaluation of Interactive Map Comparison Techniques*, in "CHI '15 Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems", Seoul, South Korea, ACM, April 2015, p. 3573-3582 [*DOI :* 10.1145/2702123.2702130], https://hal.inria.fr/hal-01144163.

[6] R. Morales González, C. Appert, G. Bailly, E. Pietriga.*TouchTokens: Guiding Touch Patterns with Passive Tokens*, in "2016 CHI Conference on Human Factors in Computing Systems", San Jose, CA, United States, May 2016 [*DOI :* 10.1145/2858036.2858041], https://hal.archives-ouvertes.fr/hal-01315130.

[7] H. Olafsdottir, C. Appert.*Multi-Touch Gestures for Discrete and Continuous Control*, in "International Working Conference on Advanced Visual Interfaces (AVI)", Como, Italy, May 2014, 8 [*DOI :* 10.1145/2598153.2598169], https://hal.archives-ouvertes.fr/hal-00998971.

[8] E. PIETRIGA, F. DEL CAMPO, A. IBSEN, R. PRIMET, C. APPERT, O. CHAPUIS, M. HEMPEL, R. MUÑOZ, S. EYHERAMENDY, A. JORDAN, H. DOLE.*Exploratory Visualization of Astronomical Data on Ultra-high-resolution Wall Displays*, in "Proceedings SPIE", July 2016, vol. 9913, 15 [*DOI :* 10.1117/12.2231191], https://hal.inria.fr/hal-01350722.

[9] A. PROUZEAU, A. BEZERIANOS, O. CHAPUIS.*Evaluating Multi-User Selection for Exploring Graph Topology on Wall-Displays*, in "IEEE Transactions on Visualization and Computer Graphics", August 2017, vol. 23, n^o 8, p. 1936–1951 [*DOI :* 10.1109/TVCG.2016.2592906], https://hal.archives-ouvertes.fr/hal-01348578.

[10] T. TSANDILAS, A. BEZERIANOS, T. JACOB.*SketchSliders: Sketching Widgets for Visual Exploration on Wall Displays*, in "Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems", Seoul, South Korea, ACM, April 2015, p. 3255-3264 [*DOI :* 10.1145/2702123.2702129], https://hal.archives-ouvertes.fr/hal-01144312.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] C. APPERT.*From Direct manipulation to Gestures: Moving the Expressive Power from the Displays to the Fingers*, Paris-Sud XI, June 2017, Habilitation à diriger des recherches, https://tel.archives-ouvertes.fr/tel-01557524.

Articles in International Peer-Reviewed Journal

[12] N. BOUKHELIFA, A. BEZERIANOS, W. CANCINO, E. LUTTON.*Evolutionary Visual Exploration: Evaluation of an IEC Framework for Guided Visual Search*, in "Evolutionary Computation", 2017, to appear [*DOI :* 10.1162/EVCO_A_00161], https://hal.inria.fr/hal-01218959.

[13] A.-S. DADZIE, E. PIETRIGA.*Visualisation of Linked Data - Reprise*, in "Open Journal Of Semantic Web", 2017, vol. 8, n^o 1, p. 1 - 21 [*DOI :* 10.3233/SW-160249], https://hal.inria.fr/hal-01406437.

[14] E. DIMARA, A. BEZERIANOS, P. DRAGICEVIC.*The Attraction Effect in Information Visualization*, in "IEEE Transactions on Visualization and Computer Graphics", 2017, vol. 23, n^o 1 [*DOI :* 10.1109/TVCG.2016.2598594], https://hal.inria.fr/hal-01355750.

[15] E. DIMARA, A. BEZERIANOS, P. DRAGICEVIC.*Conceptual and Methodological Issues in Evaluating Multidimensional Visualizations for Decision Support*, in "IEEE Transactions on Visualization and Computer Graphics", 2018, vol. 24 [*DOI :* 10.1109/TVCG.2017.2745138], https://hal.inria.fr/hal-01584729.

[16] J. FUCHS, P. ISENBERG, A. BEZERIANOS, D. KEIM.*A Systematic Review of Experimental Studies on Data Glyphs*, in "IEEE Transactions on Visualization and Computer Graphics", July 2017, vol. 23, n^o 7, p. 1863-1879 [*DOI :* 10.1109/TVCG.2016.2549018], https://hal.inria.fr/hal-01378429.

[17] M.-J. LOBO, C. APPERT, E. PIETRIGA.*MapMosaic: Dynamic Layer Compositing for Interactive Geovisualization*, in "International Journal of Geographical Information Science", May 2017, vol. 31, n^o 9, p. 1818 - 1845 [*DOI :* 10.1109/TVCG.2007.70570], https://hal.inria.fr/hal-01562084.

[18] A. PROUZEAU, A. BEZERIANOS, O. CHAPUIS.*Evaluating Multi-User Selection for Exploring Graph Topology on Wall-Displays*, in "IEEE Transactions on Visualization and Computer Graphics", August 2017,

vol. 23, n^{o} 8, p. 1936–1951 [*DOI :* 10.1109/TVCG.2016.2592906], https://hal.archives-ouvertes.fr/hal-01348578.

International Conferences with Proceedings

[19] E. DIMARA, A. BEZERIANOS, P. DRAGICEVIC.*Narratives in Crowdsourced Evaluation of Visualizations: A Double-Edged Sword?*, in "ACM Conference on Human Factors in Computing Systems (CHI)", Denver, United States, May 2017 [*DOI :* 10.1145/3025453.3025870], https://hal.inria.fr/hal-01448095.

[20] B. FRUCHARD, E. LECOLINET, O. CHAPUIS.*MarkPad: Augmenting Touchpads for Command Selection*, in "Proceedings of the 35th international conference on Human factors in computing systems", Denver, United States, CHI '17, May 2017 [*DOI :* 10.1145/3025453.3025486], https://hal.archives-ouvertes.fr/hal-01437093.

[21] V. IVANOVA, B. BACH, E. PIETRIGA, P. LAMBRIX.*Alignment Cubes: Towards Interactive Visual Exploration and Evaluation of Multiple Ontology Alignments*, in "ISWC '17 - Proceedings of the 16th International Semantic Web Conference", Vienna, Austria, Springer, October 2017, p. 400-417 [*DOI :* 10.1007/978-3-319-68288-4_24], https://hal.inria.fr/hal-01649864.

[22] C. LIU, O. CHAPUIS, M. BEAUDOUIN-LAFON, E. LECOLINET.*CoReach: Cooperative Gestures for Data Manipulation on Wall-sized Displays*, in "Proceedings of the 35th international conference on Human factors in computing systems", Denver, United States, ACM (editor), CHI '17, May 2017 [*DOI :* 10.1145/3025453.3025594], https://hal.archives-ouvertes.fr/hal-01437091.

[23] R. MORALES GONZÁLEZ, C. APPERT, G. BAILLY, E. PIETRIGA.*Passive yet Expressive TouchTokens*, in "Proceedings of the 35th SIGCHI conference on Human Factors in computing systems", Denver, United States, ACM (editor), May 2017, p. 3741 - 3745 [*DOI :* 10.1145/1978942.1979143], https://hal.inria.fr/hal-01562021.

[24] A. PROUZEAU, A. BEZERIANOS, O. CHAPUIS.*Trade-offs Between a Vertical Shared Display and Two Desktops in a Collaborative Path-Finding Task*, in "Proceedings of Graphics Interface 2017", Edmonton, Canada, May 2017, https://hal.archives-ouvertes.fr/hal-01506678.

[25] I. SADEH, I. OYA, J. SCHWARZ, E. PIETRIGA, D. DEZMAN.*The Graphical User Interface of the Operator of the Cherenkov Telescope Array*, in "Proceedings of the 2017 ICALEPCS conference", Barcelona, Spain, October 2017, https://arxiv.org/abs/1710.07117 - Proceedings of the 2017 ICALEPCS conference [*DOI :* 10.07117], https://hal.inria.fr/hal-01650408.

References in notes

[26] M. R. ENDSLEY, D. G. JONES (editors). *Designing for Situation Awareness: an Approach to User-Centered Design*, CRC Press, Taylor & Francis, 2012.

[27] S. ABITEBOUL, P. BUNEMAN, D. SUCIU.*Data on the Web: From Relations to Semistructured Data and XML*, Morgan Kaufmann, 1999.

[28] S. ABITEBOUL, P. SENELLART, V. VIANU.*The ERC Webdam on Foundations of Web Data Management*, in "Proc. of the 21st International Conference Companion on World Wide Web", WWW '12 Companion, ACM, 2012, p. 211–214, http://doi.acm.org/10.1145/2187980.2188011.

[29] J. Alvina, C. Appert, O. Chapuis, E. Pietriga.*RouteLens: Easy Route Following for Map Applications*, in "Proceedings of the International Working Conference on Advanced Visual Interfaces", AVI '14, ACM, 2014, p. 125–128 [*DOI :* 10.1145/2598153.2598200].

[30] C. Appert, O. Chapuis, E. Pietriga.*High-precision magnification lenses*, in "Proceedings of the 28th international conference on Human factors in computing systems", CHI '10, ACM, 2010, p. 273–282 [*DOI :* 10.1145/1753326.1753366].

[31] C. Appert, J.-D. Fekete.*OrthoZoom Scroller: 1D Multi-scale Navigation*, in "Proceedings of the SIGCHI Conference on Human Factors in Computing Systems", New York, NY, USA, CHI '06, ACM, 2006, p. 21–30, http://doi.acm.org/10.1145/1124772.1124776.

[32] B. Bach, E. Pietriga, J.-D. Fekete.*Visualizing Dynamic Networks with Matrix Cubes*, in "Proceedings of the 2014 Annual Conference on Human Factors in Computing Systems (CHI 2014)", Toronto, Canada, ACM, April 2014, p. 877-886 [*DOI :* 10.1145/2556288.2557010], https://hal.inria.fr/hal-00931911.

[33] M. Beaudouin-Lafon, S. Huot, M. Nancel, W. Mackay, E. Pietriga, R. Primet, J. Wagner, O. Chapuis, C. Pillias, J. R. Eagan, T. Gjerlufsen, C. Klokmose.*Multisurface Interaction in the WILD Room*, in "IEEE Computer", 2012, vol. 45, n^o^ 4, p. 48-56 [*DOI :* 10.1109/MC.2012.110].

[34] E. Benson, D. Karger.*End-users Publishing Structured Information on the Web: An Observational Study of What, Why, and How*, in "Proc. of the SIGCHI Conference on Human Factors in Computing Systems", CHI '14, ACM, 2014, p. 1265–1274, http://doi.acm.org/10.1145/2556288.2557036.

[35] J. Berndtsson, M. Normark.*The coordinative functions of flight strips: air traffic control work revisited*, in "Proceedings of the international ACM SIGGROUP conference on Supporting group work", GROUP '99, ACM, 1999, p. 101–110, http://doi.acm.org/10.1145/320297.320308.

[36] T. Berners-Lee, Y. Chen, L. Chilton, D. Connolly, R. Dhanaraj, J. Hollenbach, A. Lerer, D. Sheets.*Tabulator: Exploring and Analyzing linked data on the Semantic Web*, in "Proc. of the International Semantic Web User Interaction Workshop", 2006.

[37] J. Bertin.*Semiology of Graphics*, University of Wisconsin Press, 1983.

[38] A. Bezerianos, R. Balakrishnan.*View and Space Management on Large Displays*, in "IEEE Computer Graphics and Applications", 2005, vol. 25, n^o^ 4, p. 34-43.

[39] A. Bezerianos.*Using alternative views for layout, comparison and context switching tasks in wall displays*, in "OZCHI", 2007, p. 303-310.

[40] D. Bonnet, C. Appert, M. Beaudouin-Lafon.*Extending the Vocabulary of Touch Events with ThumbRock*, in "Proceedings of Graphics Interface", GI '13, CIPS, 2013, p. 221-228, http://dl.acm.org/citation.cfm?id=2532129.2532166.

[41] N. Boukhelifa, A. Bezerianos, W. Cancino, E. Lutton.*Evolutionary Visual Exploration: Evaluation of an IEC Framework for Guided Visual Search*, in "Evolutionary Computation", 2017, to appear [*DOI :* 10.1162/EVCO_a_00161], https://hal.inria.fr/hal-01218959.

[42] A. Collins, A. Bezerianos, G. McEwan, M. Rittenbruch, R. Wasinger, J. Kay.*Understanding File Access Mechanisms for Embedded Ubicomp Collaboration Interfaces*, in "Proceedings of the 11th International Conference on Ubiquitous Computing", New York, NY, USA, UbiComp '09, ACM, 2009, p. 135–144, http://doi.acm.org/10.1145/1620545.1620567.

[43] A.-S. Dadzie, M. Rowe.*Approaches to Visualising Linked Data: A Survey*, in "Semantic Web", 2011, vol. 2, n° 2, http://dx.doi.org/10.3233/SW-2011-0037.

[44] A. Dix, G. Lepouras, A. Katifori.*From the web of data to a world of action*, in "Web Semantics", 2010, vol. 8, p. 394–408, http://dx.doi.org/10.1016/j.websem.2010.04.007.

[45] D. Fensel, J. Hendler, H. Lieberman, W. Wahlster.*Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*, MIT Press, 2005.

[46] C. Forlines, A. Esenther, C. Shen, D. Wigdor, K. Ryall.*Multi-user, Multi-display Interaction with a Single-user, Single-display Geospatial Application*, in "Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology", UIST '06, ACM, 2006, p. 273–276, http://doi.acm.org/10.1145/1166253.1166296.

[47] J. Grudin.*Computer-Supported Cooperative Work: History and Focus*, in "IEEE Computer", May 1994, vol. 27, n° 5, p. 19–26, http://dx.doi.org/10.1109/2.291294.

[48] T. Heath, C. Bizer.*Linked Data: Evolving the Web into a Global Data Space*, Morgan & Claypool, 2011.

[49] J. Hebeler, M. Fisher, R. Blace, A. Perez-Lopez.*Semantic web programming*, Wiley & Sons, 2011.

[50] M. Hildebrand, J. van Ossenbruggen, L. Hardman.*/facet: A Browser for Heterogeneous Semantic Web Repositories*, in "ISWC'06: Proc. of the International Semantic Web Conference", Springer, 2006, p. 272–285.

[51] D. F. Huynh, R. C. Miller, D. Karger.*Potluck: Data Mash-up Tool for Casual Users*, in "Web Semantics", 2008, vol. 6, n° 4, p. 274–282, http://dx.doi.org/10.1016/j.websem.2008.09.005.

[52] H. Ishii, B. Ullmer.*Tangible Bits: Towards Seamless Interfaces Between People, Bits and Atoms*, in "Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems", CHI '97, ACM, 1997, p. 234–241, http://doi.acm.org/10.1145/258549.258715.

[53] J. A. Jacko.*Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications, 3rd Edition*, 3rd, CRC Press, 2012.

[54] D. Karger, S. Ostler, R. Lee.*The Web Page As a WYSIWYG End-user Customizable Database-backed Information Management Application*, in "Proc. of the Symposium on User Interface Software and Technology", UIST '09, ACM, 2009, p. 257–260, http://doi.acm.org/10.1145/1622176.1622223.

[55] A. Katifori, C. Halatsis, G. Lepouras, C. Vassilakis, E. Giannopoulou.*Ontology visualization methods—a survey*, in "ACM Computing Surveys", 2007, vol. 39, n° 4, p. 10:1-10:42.

[56] C. LIU, O. CHAPUIS, M. BEAUDOUIN-LAFON, E. LECOLINET.*Shared Interaction on a Wall-Sized Display in a Data Manipulation Task*, in "Proceedings of the 34th international conference on Human factors in computing systems", San Jose, United States, ACM (editor), CHI '16, SigCHI and ACM, May 2016, p. 1-12 [*DOI :* 10.1145/2858036.2858039], https://hal.inria.fr/hal-01275535.

[57] C. LIU, O. CHAPUIS, M. BEAUDOUIN-LAFON, E. LECOLINET, W. MACKAY.*Effects of Display Size and Navigation Type on a Classification Task*, in "Proceedings of the 32nd international conference on Human factors in computing systems", CHI '14, ACM, 2014, p. 4147-4156, http://doi.acm.org/10.1145/2556288.2557020.

[58] W. MACKAY.*Is paper safer? The role of paper flight strips in air traffic control*, in "ACM Trans. Comput.-Hum. Interact.", December 1999, vol. 6, p. 311–340, http://doi.acm.org/10.1145/331490.331491.

[59] R. MORALES, C. APPERT, G. BAILLY, E. PIETRIGA.*TouchTokens: Guiding Touch Patterns with Passive Tokens*, in "Proceedings of the SIGCHI Conference on Human Factors in Computing Systems", New York, NY, USA, CHI '16, ACM, 2016.

[60] T. MUNZNER.*Visualization Analysis and Design*, CRC Press, 2014.

[61] M. NANCEL, O. CHAPUIS, E. PIETRIGA, X.-D. YANG, P. P. IRANI, M. BEAUDOUIN-LAFON.*High-precision pointing on large wall displays using small handheld devices*, in "Proceedings of the SIGCHI Conference on Human Factors in Computing Systems", CHI '13, ACM, 2013, p. 831–840, http://doi.acm.org/10.1145/2470654.2470773.

[62] M. NANCEL, E. PIETRIGA, O. CHAPUIS, M. BEAUDOUIN-LAFON.*Mid-air Pointing on Ultra-Walls*, in "ACM Transactions on Computer-Human Interaction", August 2015, vol. 22, n° 5, p. 21:1–21:62 [*DOI :* 10.1145/2766448], https://hal.inria.fr/hal-01184544.

[63] A.-C. NGONGA NGOMO, L. BÜHMANN, C. UNGER, J. LEHMANN, D. GERBER.*Sorry, I Don'T Speak SPARQL: Translating SPARQL Queries into Natural Language*, in "Proc. of the International Conference on World Wide Web", WWW '13, ACM, 2013, p. 977–988, http://dl.acm.org/citation.cfm?id=2488388.2488473.

[64] H. OLAFSDOTTIR, T. TSANDILAS, C. APPERT.*Prospective Motor Control on Tabletops: Planning Grasp for Multitouch Interaction*, in "Proceedings of the 32nd international conference on Human factors in computing systems", CHI '14, ACM, 2014, p. 2893-2902, http://doi.acm.org/10.1145/2556288.2557029.

[65] E. PIETRIGA, C. APPERT, M. BEAUDOUIN-LAFON.*Pointing and beyond: an operationalization and preliminary evaluation of multi-scale searching*, in "CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems", ACM Press, 2007, p. 1215–1224 [*DOI :* 10.1145/1240624.1240808].

[66] E. PIETRIGA, C. APPERT.*Sigma lenses: focus-context transitions combining space, time and translucence*, in "CHI '08: Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems", ACM, 2008, p. 1343–1352 [*DOI :* 10.1145/1357054.1357264].

[67] E. PIETRIGA, O. BAU, C. APPERT.*Representation-Independent In-Place Magnification with Sigma Lenses*, in "IEEE Transactions on Visualization and Computer Graphics (TVCG)", 2010, vol. 16, n° 03, p. 455-467 [*DOI :* 10.1109/TVCG.2009.98].

[68] E. PIETRIGA, C. BIZER, D. KARGER, R. LEE.*Fresnel - A Browser-Independent Presentation Vocabulary for RDF*, in "Proceedings of the 5th International Semantic Web Conference (ISWC 2006)", Springer, November 2006, p. 158-171.

[69] E. PIETRIGA, S. HUOT, M. NANCEL, R. PRIMET.*Rapid development of user interfaces on cluster-driven wall displays with jBricks*, in "Proceedings of the 3rd ACM SIGCHI symposium on Engineering interactive computing systems", EICS '11, ACM, 2011, p. 185–190, http://doi.acm.org/10.1145/1996461.1996518.

[70] E. PIETRIGA.*A Toolkit for Addressing HCI Issues in Visual Language Environments*, in "IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)", IEEE Computer Society, 2005, p. 145-152 [*DOI :* 10.1109/VLHCC.2005.11].

[71] C. PINDAT, E. PIETRIGA, O. CHAPUIS, C. PUECH.*JellyLens: Content-Aware Adaptive Lenses*, in "UIST - 25th Symposium on User Interface Software and Technology - 2012", Cambridge, MA, United States, Proceedings of the 25th Symposium on User Interface Software and Technology, ACM, October 2012, p. 261-270 [*DOI :* 10.1145/2380116.2380150], https://hal.inria.fr/hal-00721574.

[72] C. PINDAT, E. PIETRIGA, O. CHAPUIS, C. PUECH.*Drilling into complex 3D models with gimlenses*, in "Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology", VRST '13, ACM, 2013, p. 223–230, http://doi.acm.org/10.1145/2503713.2503714.

[73] N. SHADBOLT, T. BERNERS-LEE, W. HALL.*The Semantic Web Revisited*, in "IEEE Intelligent Systems", 2006, vol. 21, n^{o} 3, p. 96-101, http://doi.ieeecomputersociety.org/10.1109/MIS.2006.62.

[74] O. SHAER, E. HORNECKER.*Tangible User Interfaces: Past, Present, and Future Directions*, in "Found. Trends Hum.-Comput. Interact.", January 2010, vol. 3, n^{o} 1–2, p. 1–137, http://dx.doi.org/10.1561/1100000026.

[75] B. SHNEIDERMAN, C. PLAISANT.*Designing the User Interface, 4th Edition*, Pearson Addison Wesley, 2005.

[76] D. SPELMEZAN, C. APPERT, O. CHAPUIS, E. PIETRIGA.*Controlling widgets with one power-up button*, in "Proceedings of the 26th annual ACM symposium on User interface software and technology", UIST '13, ACM, 2013, p. 71–74, http://doi.acm.org/10.1145/2501988.2502025.

[77] D. SPELMEZAN, C. APPERT, O. CHAPUIS, E. PIETRIGA.*Side pressure for bidirectional navigation on small devices*, in "Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services", MobileHCI '13, ACM, 2013, p. 11–20, http://doi.acm.org/10.1145/2493190.2493199.

[78] G. TOUYA, J.-F. GIRRES.*ScaleMaster 2.0: a ScaleMaster extension to monitor automatic multi-scales generalizations*, in "Cartography and Geographic Information Science", 2013, vol. 40, n^{o} 3, p. 192–200, http://dx.doi.org/10.1080/15230406.2013.809233.

[79] G. TOUYA.*Social Welfare to Assess the Global Legibility of a Generalized Map*, in "International Conference on Geographic Information", Springer, 2012, p. 198–211, http://dx.doi.org/10.1007/978-3-642-33024-7_15.

[80] T. TSANDILAS, C. APPERT, A. BEZERIANOS, D. BONNET.*Coordination of Tilt and Touch in One- and Two-Handed Use*, in "Proceedings of the 32nd international conference on Human factors in computing systems", CHI '14, ACM, 2014, p. 2001-2004, http://doi.acm.org/10.1145/2556288.2557088.

[81] G. TUMMARELLO, R. DELBRU, E. OREN.*Sindice.com: Weaving the Open Linked Data*, in "Proc. of the 6th Int. Semantic Web Conf. (ISWC2007)", Springer-Verlag, 2007, p. 552–565.

[82] C. WARE.*Information visualization: perception for design*, Elsevier, 2012.

[83] S. ZHAI, P. O. KRISTENSSON, C. APPERT, T. H. ANDERSEN, X. CAO.*Foundational Issues in Touch-Screen Stroke Gesture Design - An Integrative Review*, in "Foundations and Trends in Human-Computer Interaction", December 2012, vol. 5, n^{o} 2, p. 97-205 [*DOI :* 10.1561/1100000012], https://hal.inria.fr/hal-00765046.

[84] R. A. DE ALMEIDA, C. PILLIAS, E. PIETRIGA, P. CUBAUD.*Looking behind bezels: french windows for wall displays*, in "Proceedings of the International Working Conference on Advanced Visual Interfaces", AVI ’12, ACM, 2012, p. 124–131, http://doi.acm.org/10.1145/2254556.2254581.

Project-Team INFINE

INFormation NEtworks

RESEARCH CENTER
Saclay - Île-de-France

THEME
Networks and Telecommunications

Table of contents

Project-Team INFINE

Creation of the Team: 2014 April 01, updated into Project-Team: 2016 December 01, end of the Project-Team: 2017 December 31

Keywords:

Computer Science and Digital Science:

A1.2.3. - Routing
A1.2.5. - Internet of things
A1.2.6. - Sensor networks
A1.2.9. - Social Networks
A1.3. - Distributed Systems
A2.6.1. - Operating systems
A3.3.2. - Data mining
A3.4.2. - Unsupervised learning
A3.5. - Social networks
A3.5.1. - Analysis of large graphs
A3.5.2. - Recommendation systems
A4.8. - Privacy-enhancing technologies
A5.9.2. - Estimation, modeling

Other Research Topics and Application Domains:

B4.4. - Energy delivery
B4.4.1. - Smart grids
B6.4. - Internet of things
B8.1.2. - Sensor networks for smart buildings
B8.2. - Connected city

1. Personnel

Research Scientists

Laurent Massoulií¿œ [Inria, Team leader, Researcher, HDR]
Cedric Adjih [Inria, Researcher]
Emmanuel Baccelli [Inria, Researcher, HDR]
Aline Carneiro Viana [Inria, Researcher, HDR]

External Collaborators

Ichrak Amdouni [Ecole Nationale d'Ingenieur de Sousse]
Sahar Hoteit [CentraleSupelec]
Marco Fiore [CNR Torino]
Kaspar Schleiser [Fu Berlin]

Technical Staff

Francisco Javier Acosta Padilla [Inria]
Benoit Formet [Inria]
Raul Fuentes Samaniego [Inria, from May 2017]
Oliver Hahm [until Mar 2017]
Virag Shah [Inria, Engineers]

PhD Students

Rafael Costa [CAPES, from Jun 2017]
Loic Dauphin [Inria]
Hirah Malik [Inria, from Nov 2017]
Roni Shigueta [CAPES]
Guangshuo Chen [Inria]
Rémi Varloot [Microsoft Research-Inria Joint Centre]
Lennart Gulikers [Microsoft Research-Inria Joint Centre]

Post-Doctoral Fellow

Ehsan Ebrahimi Khaleghi [Inria, Post-Doctoral Fellow, from May 2016 until July 2017]

Visiting Scientist

Julinda Stefa [Sapienza University of Rome, from Jun 2017 until August 2017]

Administrative Assistant

Helene Bessin Rousseau [Inria]

2. Overall Objectives

2.1. Overall Objectives

The INFINE proposal aims to design and analyse novel communication paradigms, protocols and architectures based on concepts of ultra distributed, information- and user-centric networking. The project is motivated by the recent and forthcoming evolution of Internet uses. Based on an information- and user-centric perspective, not only does it address issues pertaining to physical communication networks such as traffic routing, regulation and caching, but also issues about online social networks such as content recommendation and privacy protection.

INFINE team is engaged in research along three main themes: *Online social networking*, *Traffic and Resource Management*, and *Spontaneous Wireless Networks*. All these research activities encompass both theoretical research (on elaboration of models, algorithms, protocols and formal characterization of their performance), and applied research (to validate and/or experiment the proposed concepts against real networking scenarios). INFINE fits in the theme ”Networks and Telecommunications” of the research field ”Networks, Systems and Services, Distributed Computing” at Inria.

2.2. New challenging demands

Nowadays, we use networks not only to transport information from where it resides to ourselves but also, with online social networks, to determine what information might be of interest to us. Such a social recommendation functionality holds the promise of allowing us to access more relevant information. At the same time there is ample scope for improving its efficiency. Moreover it creates threats to user privacy.

At the same time, the physical context in which we access communication networks has drastically changed. While in the past, Internet was mostly accessed through fixed desktop computers, users are now mobile about 50% of their time online. In addition, while communicating machines used to be sparse and wired, with the advent of the Internet of Things we now evolve in a dense, interconnected environment of heterogeneous devices communicating via wireless and/or via wires.

This new context of Internet uses challenges several aspects of currently deployed networks. Some aspects pertain to the physical architecture of the Internet. In particular, at the core of the Internet, a drastic increase in volume of data traffic is anticipated due to the emergence of new applications, generalization of cloud services, or the advent of the Internet of Things (IoT) and Machine-to-Machine (M2M) communications. On the other hand, at the edge of the Internet, user mobility and today’s pervasiveness of computing devices with increasingly higher capabilities (i.e., processing, storage, sensing) have a fundamental impact on the adequacy of algorithms and communication mechanisms.

Other aspects concern the logical architecture of the network. For instance, currently deployed protocols at layers above IP must now carry massive publish-subscribe traffic, preserve user privacy, be social-aware, and support delay tolerant communications and paradigms for which they were not initially designed. Concerning actual content distribution, the avalanche of data and privacy concerns puts more and more pressure on filter/push mechanisms to provide users with relevant information.

While considering physical and logical aspects of networks, the INFINE team will pursue research activities combining theoretical and experimental approaches.

2.3. Research agenda

Our general goal is to develop distributed mechanisms for optimizing the operation of networks both at the mentioned logical and physical levels of the architecture. Taking an information- and user-centric perspective, we envision networks as means to convey relevant information to users, while adapting to customary practices (in terms of context, interests, or content demands) of such users. At the logical level, online social networks (OSNs) allow users to choose what information to access. At the physical level, communication, computation, and memory resources allow users to retrieve some content eventually selected on the basis of the online social network.

The two setups feature scarce resources: for instance, in OSNs, these are the users' budget of attention, which must be used sparingly by recommending only relatively few potential content items. At the physical level this is typically the channels' capacity or networking resources, which cannot be oversubscribed.

Beyond a formal resemblance between the optimizations that one must carry at these two levels, there is a strong commonality in the methods adequate for conducting optimizations in the two setups. To illustrate this point, consider *contact recommendation*, that is a key objective in our agenda on online social networks. This entails automatically proposing to users potential contacts for optimizing the subsequent efficiency of social content filtering. We envision addressing contact recommendation by first performing some community detection, i.e. identification of similarly behaving users. Similarly, at the physical level, user-centric approaches, sometimes also related to community detection, have guided routing decisions in challenged network environments, where delay-tolerant networking is used. Still, associated with dynamic centrality metrics, community detection can guide the replication of a specific content in well-selected users, while exploiting the advantages of distributed decentralized storage and opportunistic communications.

As an additional example at the logical level, we consider *content recommendation*, whereby a list of potential contents is filtered before being presented to a user, with the aim of maximizing the chance this user finds an item of interest therein. This has an exact analogue at the physical level, where by taking an information- and user-centric approach, we intend to off-load communication resources via pre-loaded content replicas at various storage points in the network. The problem of determining which content to cache so as to maximize the chance of it being accessed in the vicinity of the corresponding cache memory corresponds precisely to the aforementioned content recommendation problem.

We now detail further our agenda along three main specific axes, namely Online Social Networks, Traffic and Resource Management, and Internet of Things/Spontaneous Wireless Networks, bearing in mind that we will develop generic solutions relevant to several of these axes wherever possible.

3. Research Program

3.1. Online Social Networks (OSN)

Large-scale online social networks such as Twitter or FaceBook provide a powerful means of selecting information. They rely on "social filtering", whereby pieces of information are collectively evaluated and sorted by users. This gives rise to information cascades when one item reaches a large population after spreading much like an epidemics from user to user in a viral manner. Nevertheless, such OSNs expose their users to a large amount of content of no interest to them, a sign of poor "precision" according to the terminology of information retrieval. At the same time, many more relevant content items never reach those users most interested in them. In other words, OSNs also suffer from poor "recall" performance.

This leads to a first challenge: *what determines the optimal trade-off between precision and recall in OSNs? And what mechanisms should be deployed in order to approach such an optimal trade-off?* We intend to study this question at a theoretical level, by elaborating models and analyses of social filtering, and to validate the resulting hypotheses and designs through experimentation and processing of data traces. More specifically, we envision to reach this general objective by solving the following problems.

3.1.1. Community Detection

Identification of implicit communities of like-minded users and contact recommendation for helping users "rewire" the information network for better performance. Potential schemes may include variants of spectral clustering and belief propagation-style message passing. Limitations / relative merits of candidate schemes, their robustness to noise in the input data, will be investigated.

3.1.2. Incentivization

Design of incentive mechanisms to limit the impact of users' selfishness on system behavior: efficiency should be maintained even when users are gaming the system to try and increase their estimated expertise. By offering rewards to users on the basis of their involvement in filtering and propagation of content, one might encourage them to adjust their action and contribute to increase the overall efficiency of the OSN as a content access platform.

One promising direction will be to leverage the general class of Vickrey-Clarke-Groves incentive-compatible mechanisms of economic theory to design so-called marginal utility reward mechanisms for OSN users.

3.1.3. Social Recommendation and Privacy

So far we have only alluded to the potential benefits of OSNs in terms of better information access. We now turn to the risks they create. Privacy breaches constitute the greatest of these risks: OSN users disclose a wealth of personal information and thereby expose themselves to discrimination by potential employers, insurers, lenders, government agencies etc. Such privacy concerns are not specific to OSNs: internauts' online activity is discretely tracked by companies such as Bluekai, and subsequently monetized to advertisers seeking better ad targeting. While disclosure of personal data creates a privacy risk, on the other hand it fuels personalized services and thereby potentially benefits everyone.

One line of research will be to focus on the specific application scenario of content categorization, and to characterize analytically the trade-off between user privacy protection (captured by differential privacy), accuracy of content categorization, and sample complexity (measured in number of probed users).

3.2. Traffic and Resource Management

Despite the massive increases in transmission capacity of the last few years, one has every reason to believe that networks will remain durably congested, driven among other factors by the steadily increasing demand for video content, the proliferation of smart devices (i.e., smartphones or laptops with mobile data cards), and the forecasted additional traffic due to machine-to-machine (M2M) communications. Despite this rapid traffic growth, there is still a rather limited understanding of the features protocols have to support, the characteristics of the traffic being carried and the context where it is generated. There is thus a strong need for smart protocols that transport requested information at the cheapest possible cost on the network as well as provide good quality of service to network subscribers. One particularly new aspect of up-and-coming networks is that networks are now used to not only (i) access information, but also (ii) distributively process information, en-route.

We intend to study these issues at the theoretical and protocol design levels, by elaborating models and analysis of content demands and/or mobility of network subscribers. The resulting hypothesis and designs will be validated through experimentation, simulation, or data trace processing. It is also worth mentioning the provided solutions may bring benefits to different entities in the network: to content owners (if applied at the core of Internet) or to subscribers or network operators (if applied at the edge of the Internet).

3.2.1. At the Internet Core

One important optimization variable consists in content replication: users can access the closest replica of the content they are interested in. Thus the memory resource can be used to create more replicas and reduce the usage of the bandwidth resource. Another interesting arbitrage between resources arises because content is no longer static but rather dynamic. Here are two simple examples: i) a video could be encoded at several resolutions. There is then a choice between pre-recording all possible resolutions, or alternatively synthesizing a lower-resolution version on the fly from a higher resolution version when a request arises. ii) A user requests the result of a calculation, say the average temperature in a building; this can either be kept in memory, or recomputed each time such a query arises. Optimizing the joint use of all three resources, namely bandwidth, memory, computation, is a complex task. Content Delivery Network companies such as Akamai or Limelight have worked on the memory/bandwidth trade-off for some years, but as we will explain more can be done on this. On the other hand optimizing the memory/computation trade-off has received far less attention. We aim to characterize the best possible content replication strategies by leveraging fine-grained prediction of i) users' future requests, and ii) wireless channels' future bandwidth fluctuations. In the past these two determining inputs have only been considered at a coarse-grained, aggregate level. It is important to assess how much bandwidth saving can be had by conducting finer-grained prediction. We are developing light-weight protocols for conducting these predictions and automatically instantiating the corresponding optimal replication policies. We are also investigating generic protocols for automatically trading replication for computation, focusing initially on the above video transcoding scenario.

3.2.2. At the Internet Edge

Cellular and wireless data networks are increasingly relied upon to provide users with Internet access on devices such as smartphones, laptops or tablets. In particular, the proliferation of handheld devices equipped with multiple advanced capabilities (e.g., significant CPU and memory capacities, cameras, voice to text, text to voice, GPS, sensors, wireless communication) has catalyzed a fundamental change in the way people are connected, communicate, generate and exchange data. In this evolving network environment, users' social relations, opportunistic resource availability, and proximity between users' devices are significantly shaping the use and design of future networking protocols.

One consequence of these changes is that mobile data traffic has recently experienced a staggering growth in volume: Cisco has recently foreseen that the mobile data traffic will increase 18-fold within 2016, in front of a mere 9-fold increase in connection speeds. Hence, one can observe today that the inherently centralized and terminal-centric communication paradigm of currently deployed cellular networks cannot cope with the increased traffic demand generated by smartphone users. This mismatch is likely to last because (1) forecasted mobile data traffic demand outgrows the capabilities of planned cellular technological advances such as 4G or LTE, and (2) there is strong skepticism about possible further improvements brought by 5G technology.

Congestion at the Internet's edge is thus here to stay. Solutions to this problem relates to: densify the infrastructure, opportunistically forward data among neighbors wireless devices, to offload data to alternate networks, or to bring content from the Internet closer to the subscribers. Our recent work on leveraging user mobility patterns, contact and inter-contact patterns, or content demand patterns constitute a starting point to these challenges. The projected increase of mobile data traffic demand pushes towards additional complementary offloading methods. Novel mechanisms are thus needed, which must fit both the new context that Internet users experience now, and their forecasted demands. In this realm, we will focus on new approaches leveraging ultra-distributed, user-centric approaches over IP.

3.3. Internet of Things (IoT) and Spontaneous Wireless Networks (SWN)

The unavailability of end-to-end connectivity in emergent wireless mobile networks is extremely disruptive for IP protocols. In fact, even in simpler cases of spontaneous wireless networks where end-to-end connectivity exists, such networks are still disruptive for the standard IP protocol stack, as many protocols rely on atomic link-local services (such as link-local multicast/broadcast), while these services are inherently unavailable in such networks due to their opportunistic, wireless multi hop nature. In this domain, we will aim to characterize the achievable performance in such IP-disruptive networks and to actively contribute to the design of new, deployable IP protocols that can tolerate these disruptions, while performing well enough compared to what is achievable and remaining interoperable with the rest of the Internet.

Spontaneous wireless networking is also a key aspect of the Internet of Things (IoT). The IoT is indeed expected to massively use this networking paradigm to gradually connect billions of new devices to the Internet, and drastically increase communication without human source or destination – to the point where the amount of such communications will dwarf communications involving humans. Large scale user environment automation require communication protocols optimized to efficiently leverage the heterogeneous and unreliable wireless vicinity (the scope of which may vary according to the application). In fact, extreme constraints in terms of cost, CPU, battery and memory capacities are typically experienced on a substantial fraction of IoT devices. We expect that such constraints will not vanish any time soon for two reasons. On one hand the progress made over the last decade concerning the cost/performance ratio for such small devices is quite disappointing. On the other hand, the ultimate goal of the IoT is ubiquitous Internet connectivity between devices as tiny as dust particles. These constraints actually require to redesign not only the network protocol stack running on these devices, but also the software platform powering these machines. In this context, we will aim at contributing to the design of novel network protocols and software platforms optimized to fit these constraints while remaining compatible with legacy Internet.

3.3.1. Design & Development of Open Experimental IoT Platforms

Manufacturers announce on a regular basis the availability of novel tiny devices, most of them featuring network interfaces: the Internet of Things (IoT) is already here, from the hardware perspective, and it is expected in the near future that we will see a massive increase of the number of muti-purpose smart objects (from tiny sensors in industrial automation to devices like smart watches and tablets). Thus, one of the challenges is to be able to test architectures, protocols and applications, in realistic conditions and at large scale.

One necessity for research in this domain is to establish and improve IoT hardware platforms and testbeds, that integrate representative scenarios (such as Smart Energy, Home Automation etc.) and follow the evolution of technology, including radio technologies, and associated experimentation tools. For that, we plan to build upon the IoT-LAB federated testbeds, that we have participated in designing and deploying recently. We plan to further develop IoT-LAB with more heterogeneous, up-to-date IoT hardware and radios that will provide a usable and realistic experimentation environment. The goal is to provide a tool that enables testing a validation of upcoming software platforms and network stacks targeting concrete IoT deployments.

In parallel, on the software side, IoT hardware available so far made it uneasy for developers to build apps that run across heterogeneous hardware platforms. For instance Linux does not scale down to small, energy-constrained devices, while microcontroller-based OS alternatives were so far rudimentary and yield a steep learning curve and lengthy development life-cycles because they do not support standard programming and debugging tools. As a result, another necessity for research in this domain is to allow the emergence of it more powerful, unifying IOT software platforms, to bridge this gap. For that, we plan to build upon RIOT, a new open source software platform which provides a portable, Linux-like API for heterogeneous IoT hardware. We plan to continue to develop the systems and network stacks aspects of RIOT, within the open source developer community currently emerging around RIOT, which we co-founded together with Freie Universitaet Berlin. The key challenge is to improve usability and add functionalities, while maintaining architectural consistency and a small enough memory footprint. The goal is to provide an IoT software platform that can be used like Linux is used for less constrained machines, both (i) in the context of research and/or teaching, as well as (ii) in industrial contexts. Of course, we plan to use it ourselves for our own experimental research activities in the domain of IoT e.g., as an API to implement novel network protocols running on IoT hardware, to be tested and validated on IoT-LAB testbeds.

3.3.2. *Design & Standardization of Architectures and Efficient Protocols for Internet of Things*

As described before, and by definition, the Internet of Things will integrate not only a massive number of homogeneous devices (e.g., networks of wireless sensors), but also heterogeneous devices using various communication technologies. Most devices will be very constrained resources (memory resources, computational resources, energy). Communicating with (and amongst) such devices is a key challenge that we will focus on. The ability to communicate efficiently, to communicate reliably, or even just to be able to communicate at all, is non-trivial in many IoT scenarios: in this respect, we intend to develop innovative protocols, while following and contributing to standardization in this area. We will focus and base most of our work on standards developed in the context of the IETF, in working groups such as 6lo, CORE, LWIG etc., as well as IRTF research groups such as NWCRG on network coding and ICNRG on Information Centric Networking. We note however that this task goes far beyond protocol design: recently, radical rearchitecturing of the networks with new paradigms such as Information Centric Networking, ICN, (or even in wired networks, software-defined networks), have opened exciting new avenues. One of our direction of research will be to explore these content-centric approaches, and other novel architectures, in the context of IoT.

4. Highlights of the Year

4.1. Awards

Laurent Massoulié received the Grand Prix Scientifique Cino Del Duca 2017, awarded by the French Academy of Sciences for his project on *S*ocial Information Networks. More details about the price here: http://www.academie-sciences.fr/fr/Prix-en-chimie-et-sciences-du-vivant/grand-prix-scientifique-fondation-simone-et-cino-del-duca.html. More details about his project here: https://www.inria.fr/en/centre/saclay/news/laurent-massoulie-grand-prix-cino-del-duca.

Oliver Hahm received the 1st prize of Université Paris Saclay (Prix Doctorant ED STIC 2017) for his PhD work on RIOT, supervised by Emmanuel Baccelli.

4.2. Associated team - EMBRACE

2017 was the first year of the EMBRACE Associated team. The EMBRACE (lEveraging huMan Behavior for Resource AlloCation and services orchestration modEls) team is composed by members of the INFINE and by three Brazilian teams from three different Brazilian Universities. The EMBRACE project addresses the topic of designing efficient solutions for 5G networks taking into account human behavior, uncertainty, and heterogeneity of networking resources.

More information is available here: https://team.inria.fr/infine/embrace/

4.3. RIOT Summit 2017

We successfully organized in September 2017 the second RIOT Summit in Berlin. The RIOT Summit 2017 gathered 100+ enthusiastic industrial participants, makers and academics involved in RIOT. Relevant partners such as Cisco, Fujitsu, OTA Keys, Wolf SSL, as well as a number of SMEs and startups from various places in Europe gave talks on aspects of IoT communication, use cases IoT hardware, IoT open source community aspects and concepts for future IoT software and networks, as well as hands-on sessions and tutorials. See: http://summit.riot-os.org/#speakers.

5. New Software and Platforms

5.1. Gardinet

KEYWORD: Distributed networks
FUNCTIONAL DESCRIPTION: Gardinet (previously DragonNet) is a generic framework for network coding in wireless networks. It is a initially result of the GETRF project of the Hipercom2 team.

It is based on intra-flow coding where the source divides the flow in a sequence of payloads of equal size (padding may be used). The design keys of DragonNet are simplicity and universality, DragonNet does not use explicit or implicit knowledge about the topology (such as the direction or distance to the source, the loss rate of the links, ...). Hence, it is perfectly suited to the most dynamic wireless networks. The protocol is distributed and requires minimal coordination. DragonNet architecture is modular, it is based on 5 building blocks (LIB, SIG, Protocol, SEW and DRAGON). Each block is almost independent. This makes DragonNet generic and hence adaptable to many application scenarios. DragonNet derives from a prior protocol called DRAGONCAST. Indeed, DragonNet shares the same principles and theoretical overview of DRAGONCAST. It enriches DRAGONCAST by the information base and signaling required to perform broadcast in wireless networks and in wireless sensor networks in particular.

- Participants: Antonia Masucci, Cédric Adjih, Hana Baccouch and Ichrak Amdouni
- Contact: Cédric Adjih
- URL: http://gitlab.inria.fr/gardinet

5.2. MACACO

Mobile context-Adaptive CAching for COntent-centric networking
FUNCTIONAL DESCRIPTION: MACACOapp is developped in the context of the EU CHIST-ERA MACACO project. It consists in a mobile phone application that periodically samples phone's information on the mobility (through, e.g., GPS sensor, accelerometer and WiFi/Bluetooth/Cellular environment, connectivity type) and on the data traffic it generates (through, e.g., Internet browser history and applications data consumption). The information collected will be time-stamped and will be periodically sent to the central servers for analysis and visualization. We expect that (1) the collected information will allow us studying the correlation between mobility and content demand patterns and that (2) the results of this analysis will allow us inferring the best times and places to transfer content from/to users' phones location and/or from/to the wireless infrastructure closest to the users' phones location. Users will be also invited to fill a non-mandatory questionnaire relevant to this study. Our questionnaire collects information about the personality traits and application preferences of people. We expect that the information collected from questionnaire will allow us to analyse the correlation between users personality traits and their application preferences and interests. User's application preferences and interests will be inferred from the Internet browsing history and running app information obtained from the MACACO App.

- Participants: Aline Carneiro Viana, Katia Jaffres and Marco Fiore
- Contact: Aline Carneiro Viana
- URL: https://macaco.inria.fr/macacoapp/

5.3. RIOT

KEYWORDS: Internet of things - Wireless Sensor Networks - Iot - Sensors - Operating system - Internet protocols
SCIENTIFIC DESCRIPTION: While requiring as low as 1,5kB of RAM and 5kB or ROM, RIOT offers real time and energy efficiency capabilities, as well as a single API (partially POSIX compliant) across heterogeneous 8-bit, 16-bit and 32-bit low-hardware. This API is developer-friendly in that it enables multi-threading, standard C and C++ application programming and the use of standard debugging tools (which was not possible so far for embedded programming). On top of this, RIOT includes several network stacks, such as a standard IPv6/6LoWPAN stack and a information-centric network stack (based on CCN).

FUNCTIONAL DESCRIPTION: RIOT is an Open Source operating system that provides standard protocols for embedded systems. RIOT allows, for example, the development of applications that collect sensor data and transmit it to a central node (e.g. a server). This data can then be used for smart energy management for instance.

RIOT is specially designed for embedded systems, which are strongly constrained in memory and energy. Further, RIOT can easily be ported to different hardware devices and follows the latest evolution of IP standards.

RIOT applications can readily be tested in the FIT IoT-Lab, which provides a large-scale infrastructure facility with 3000 nodes for testing remotely small wireless devices.

- Participants: Emmanuel Baccelli and Oliver Hahm
- Partner: Freie Universität Berlin
- Contact: Emmanuel Baccelli
- URL: http://www.riot-os.org

6. New Results

6.1. Online Social Networks (OSN)

6.1.1. Capacity of Information Processing Systems

- Participants: Laurent Massoulié and Kuang Xu

We propose and analyze a family of information processing systems, where a finite set of experts or servers are employed to extract information about a stream of incoming jobs. Each job is associated with a hidden label drawn from some prior distribution. An inspection by an expert produces a noisy outcome that depends both on the job's hidden label and the type of the expert, and occupies the expert for a finite time duration. A decision maker's task is to dynamically assign inspections so that the resulting outcomes can be used to accurately recover the labels of all jobs, while keeping the system stable. Among our chief motivations are applications in crowd-sourcing, diagnostics, and experiment designs, where one wishes to efficiently learn the nature of a large number of items, using a finite pool of computational resources or human agents. We focus on the capacity of such an information processing system. Given a level of accuracy guarantee, we ask how many experts are needed in order to stabilize the system, and through what inspection architecture. Our main result provides an adaptive inspection policy that is asymptotically optimal in the following sense: the ratio between the required number of experts under our policy and the theoretical optimal converges to one, as the probability of error in label recovery tends to zero. This work was firstly accepted and presented at the COLT conference.

6.2. Spontaneous Wireless Networks (SWN)

6.2.1. Spatio-Temporal Prediction of Cellular Data Traffic

- Participants: Guangshuo Chen, Aline Carneiro Viana, Marco Fiore, Carlos Sarraute

The understanding of human behaviors is a central question in multi-disciplinary research and has contributed to a wide range of applications. The ability to foresee human activities has essential implications in many aspects of cellular networks. In particular, the high availability of mobility prediction can enable various application scenarios such as location-based recommendation, home automation, and location-related data dissemination; the better understanding of future mobile data traffic demand can help to improve the design of solutions for network load balancing, aiming at improving the quality of Internet-based mobile services. Although a large and growing body of literature has investigated the topic of predicting human mobility, there has been little discussion in anticipating mobile data traffic in cellular networks, especially in spatiotemporal view of individuals. We address the problem of understanding spatiotemporal mobile data traffic demand for individuals and perform an theoretical and empirical analysis of jointly predicting human whereabouts and mobile data traffic, by collaboratively mining human mobility dataset and mobile data traffic dataset. Our contributions are summarized as follows:

- We investigate the limits of predictability by measuring the maximum predictability that any algorithm has potential to achieve based on tools of information theory. Our theoretical analysis shows that it is theoretically possible to anticipate the individual demand with a typical accuracy of 75% despite the heterogeneity of users and with an improved accuracy of 80% using joint prediction with mobility information. This work was published at the IEEE LCN 2017 international conference and the Technical report RT-0483 brings a full description of the work, which is being prepared for a journal submission.
- We evaluate the state-of-the-art predictors and propose novel solutions for predicting mobile data traffic via machine learning algorithms. Our data-driven test on the performance of these predictors show that the 2nd order Markov predictor outperforms all the legacy time series predictors. It can achieve a mean accuracy of 62% but can hardly have an enhancement from knowing human mobility information. Besides, based on machine learning techniques, our proposed solutions can achieve a typical accuracy of 70% and have a 1% 5% degree of improvement by learning individual whereabouts (what confirms the predictability theoretical results). Finally, our analysis show that knowing mobile data traffic of a user can significantly help the prediction of his whereabouts for 50% of the users, leading to an improvement up to 10% regarding accuracy. The Technical Report hal-01675573 brings more details on this work. A conference paper is also in preparation.

All those works were performed in the context of the Guangshuo Chen's PhD thesis, who will defend in March 2018.

6.2.2. Human Mobility completion of Sparse Call Detail Records for Mobility Analysis

- Participants: Guangshuo Chen, Aline Carneiro Viana, Marco Fiore, Sahar Hoteit

Call Detail Records (CDR) are an important source of information in the study of diverse aspects of human mobility. The accuracy of mobility information granted by CDR strongly depends on the radio access infrastructure deployment and the frequency of interactions between mobile users and the network. As cellular network deployment is highly irregular and interaction frequencies are typically low, CDR are often characterized by spatial and temporal sparsity, which, in turn, can bias mobility analyses based on such data. In this paper, we precisely address this subject. First, we evaluate the spatial error in CDR, caused by approximating user positions with cell tower locations. Second, we assess the impact of the limited spatial and temporal granularity of CDR on the estimation of standard mobility metrics. Third, we propose novel and effective techniques to reduce temporal sparsity in CDR, by leveraging regularity in human movement patterns.

These works have been published as invited papers at the ACM CHANTS 2016 workshop (in conjunction with ACM MobiCom 2016) and at the IEEE DAWM workshop (in conjunction with IEEE Percom 2017). A journal version (also registered as TR: hal-01646608) is in revision at the Computer Communication Elsevier Journal, and got the first notification asking for minor revisions. Finally, a new completion methodology improving the previously described that leverages tensor factorization was designed and will be submitted to a journal: the technical report hal-01675570 describes this work.

6.2.3. Sampling frequency of human mobility

- Participants: Panagiota Katsikouli, Aline Carneiro Viana, Marco Fiore, Alberto Tarable

Recent studies have leveraged tracking techniques based on positioning technologies to discover new knowledge about human mobility. These investigations have revealed, among others, a high spatiotemporal regularity of individual movement patterns. Building on these findings, we aim at answering the question *"at what frequency should one sample individual human movements so that they can be reconstructed from the collected samples with minimum loss of information?"*. Our quest for a response leads to the discovery of *(i)* seemingly universal spectral properties of human mobility, and *(ii)* a linear scaling law of the localization error with respect to the sampling interval. Our findings are based on the analysis of fine-grained GPS trajectories of 119 users worldwide. The applications of our findings are related to a number of fields relevant to ubiquitous computing, such as energy-efficient mobile computing, location-based service operations, active probing of subscribers' positions in mobile networks and trajectory data compression. to an international conference in the next months. This work was published at the IEEE Globecom 2017 international conference.

We are improving the currently published sampling approache by incorporating human behavioral features at the sampling decisions to make it more adaptive. This is an on-going work with Panagiota Katsikouli, who spent 5 months in our team working as an internship and is currently doing a Post-Doc at the AGORA Inria team.

6.2.4. Inference of human personality from mobile phones datasets

- Participants: Adriano di Luzio, Aline Carneiro Viana, Julinda Stefa, Katia Jaffres

Personality research has enjoyed a strong resurgence over the past decade. Trait-based personality theories define personality as the traits that predict a person's behavior through learning and habits. Personality traits are relatively stable over time, differ across individuals, and most importantly, influence behavior. In psychology, the human personality has been modeled into a set of independent factors that, together, accurately describe any individual: The Five Factors Personality Model. This personality model presents the Big Five personality traits, often represented by the OCEAN acronym: Openness: appreciation for a variety of experiences; Conscientiousness: planning ahead rather than being spontaneous; Extraversion: being sociable, energetic and talkative; Agreeableness: being kind, sympathetic and happy to help; Neuroticism: inclined to worry or be vulnerable or temperamental.
This is a very recently started work, where we are firstly analysing the relationship between smartphone usages (i.e., social interactions, content interest, mobility, and communication) and personality traits in the Big Five Model. Most of the studies on personality traits were performed by social scientists and in particular, by psychologists. Studies reveal that one of the most distal influences shaping personality lie in the environment where development occurs. Nevertheless, the identification of precise environmental sources impacting personality is still an open research. More recently, computer science researchers have tried to extract personality from datasets collected through smartphones. Although laying the ground work to understand human personality from smartphones usage, much still remain to be investigated. Thus, we are performing analysis to study the correlation between traits and technological features. We plan then to establish a methodology to infer traints from features and consequently, to investigate how different traits influence different features.

This is an on-going work with Adriano di Luzio, who spent 4 months in our team working as an internship, and Julinda Stefa, an invited research visitor at Infine.

6.2.5. Predicting new places to visit in human mobility decision

- Participants: Maria Astefanoaei, Aline Carneiro Viana, Rik Sarkar

Most location prediction methods need a large user mobility history to accurately predict the next location of a user (markov chains, rnn). These methods are particularly good for predicting locations that are frequently visited by users, but not as good for predicting new places or how a user?s trajectories change in case of random events. We amend this by using contextual information to manage new places and random events and the movement patterns of users who exhibit similar behaviours. In this context, we plan to use the user?s profile and social ties to identify the most probable next category of locations (type of actions: entertainment, social, food etc.). Then, use subtrajectory similarity to predict the route taken to the identified area. This is an on-going work with the intern Maria Astefanoaei and her advisor, who spent 5 months in our team.

6.2.6. Data offloading decision via mobile crowdsensing

- Participants: Emanuel Lima, Aline Carneiro Viana, Ana Aguiar

With the steady growth of smart-phones sales [1], the demand for services that generate mobile data traffic has grown tremendously. WiFi offloading has been considered as a promising solution to the recent boost up of mobile data consumption that is making excessive demands on cellular networks in metropolitan areas. The idea consists in shifting the traffic off of cellular networks to WiFi networks. Characterizing the capacity and availability of a chaotic deployed dense WiFi network is crucial to understand and decide where and when to o oad data. This is the first goal of this work, where the MACACO dataset was considered in the characterization. Our final goal is the design of a decision strategy allowing a mobile phone of a user to decide if offload or not her traffic, i.e., when, where (using what Access Point in her usual mobility) and how (if the traffic will be offloadied to one or more Access Points). This is an on-going work with the intern emanuel Lima and his advisor, who spent 4 months in our team.

6.2.7. Infering friends in the crowd in Device-to-Device communication

- Participants: Rafael Costa, Aline Carneiro Viana, Leobino Sampaio, Artur Ziviani

The next generation of mobile phone networks (5G) will have to deal with spectrum bottleneck and other major challenges to serve more users with high-demanding requirements. Among those are higher scalability and data rates, lower latencies and energy consumption plus reliable ubiquitous connectivity. Thus, there is a need for a better spectrum reuse and data offloading in cellular networks while meeting user expectations. According to literature, one of the 10 key enabling technologies for 5G is device-to-device (D2D) communications, an approach based on direct user involvement. Nowadays, mobile devices are attached to human daily life activities, and therefore communication architectures using context and human behavior information are promising for the future. User-centric communication arose as an alternative to increase capillarity and to offload data traffic in cellular networks through opportunistic connections among users. Although having the user as main concern, solutions in the user-centric communication/networking area still do not see the user as an individual, but as a network active element. Hence, these solutions tend to only consider user features that can be measured from the network point of view, ignoring the ones that are intrinsic from human activity (e.g., daily routines, personality traits, etc). In this work, we plan to investigate how human-aspects and behavior can be useful to leverage future device-to-device communication. This is a recently started PhD thesis subject, aiming the design of a methodology to select next-hops in a D2D communication that will be human-aware: i.e., that will consider not only available physical resoures at the mobile device of a wireless neighbor, her mobility features and restrictions but also any information allowing to infer how much sharing willing she is.

6.3. Internet of Things (IoT) and Information Centric Networking (ICN)

6.3.1. Low-power Internet of Things with NDN and Cooperative Caching

- Participants: Oliver Hahm, Emmanuel Baccelli, Thomas C. Schmidt, Matthias Wählisch, Cédric Adjih, and Laurent Massoulié

Energy efficiency is a major driving factor in the Internet of Things (IoT). In this context, an IoT approach based on Information-Centric Networking (ICN) offers prospects for low energy consumption. Indeed, ICN can provide local in-network content caching so that relevant IoT content remains available at any time while devices are in deep-sleep mode most of the time. In our paper on the subject, we evaluated NDN enhanced with CoCa, a simple side protocol we designed to exploit content names together with smart interplay between cooperative caching and power-save sleep capabilities on IoT devices. We performed extensive, large scale experiments on real hardware with IoT networks comprising of up to 240 nodes, and on an emulator with up to 1000 nodes. We have shown in practice that, with NDN+CoCa, devices can reduce energy consumption by an order of magnitude while maintaining recent IoT content availability above 90 %. We furthermore provided auto-configuration mechanisms enabling practical ICN deployments on IoT networks of arbitrary size with NDN+CoCa. With such mechanisms, each device could autonomously configure names and auto-tune parameters to reduce energy consumption as demonstrated in our paper.

6.3.2. Information Centric Networking for the IoT Robotics

- Participants: Loic Dauphin, Emmanuel Baccelli, Cédric Adjih

In the near-future, humans will interact with swarms of low-cost, interconnected robots. Such robots will hence integrate the Internet of Things, and coin the term IoT robotics. Using ROS (Robot Operating System) is currently the dominant approach to implement distributed robotic software modules communicating with one another. ROS nodes can publish or subscribe to topics, which are named and typed data streams sent over the network. In our work on the subject, we presented preliminary work exploring the potential of using NDN as network primitive for ROS2 nodes (the newest version of ROS).

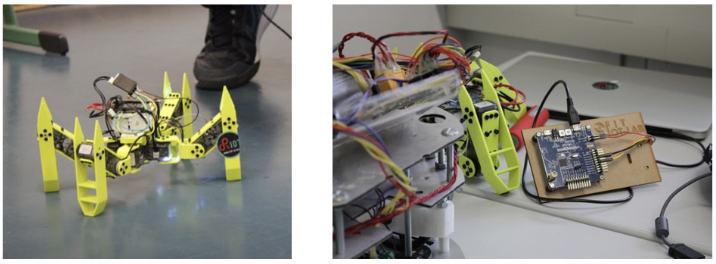


Figure 1. Demonstration of the use of NDN in IoT Robotics

6.3.3. Data Synchronization through Information Centric Networking

- Participants: Ayat Zaki Hindi, Cédric Adjih, Michel Kieffer, Claudio Weidmann

The use of Named Data Networking (NDN) for distributed multiuser applications, e.g. group messaging and file sharing, requires NDN synchronization protocols to maintain the same shared dataset (and its updates) among all nodes. ChronoSync, RoundSync, and PartialSync are some proposals to address this issue.

In our work on the subject, we focused on the state-of-the-art protocol RoundSync: we study its core features, that permit participating nodes to detect, propagate, and reconcile all changes. Particular attention is given to the case of multiple changes per round. We then proposed an improved variant, iRoundSync, that exchanges fewer messages in the multiple-change case and is more resilient to packet losses. We have quantified the performance gain of iRoundSync on a simple topology.

6.4. Internet of Things (IoT) and 5G

6.4.1. Efficient Random Access for 5G Systems: Coded Slotted Aloha

- Participants: Ehsan Ebrahimi Khaleghi, Cédric Adjih, Amira Alloum, Paul Mï¿œhlethaler, Vinod Kumar

Motivated by scenario requirements for 5G cellular networks, we have studied one of the candidate protocols for massive random access: the family of random access methods known asCoded Slotted ALOHA (CSA). A recent trend in research hasexplored aspects of such methods in various contexts, but oneaspect has not been fully taken into account: the impact of pathloss, which is a major design constraint in long-range wireless networks. In one article, we explored the behavior of CSA, by focusing on the path loss component correlated to the distanceto the base station. Path loss provides opportunities for capture,improving the performance of CSA. We revised methods for estimating CSA behavior, provide bounds of performance, and then,focusing on the achievable throughput, we extensively explored the key parameters, and their associated gain (experimentally). Our results shed light on the behavior of the optimal distribution of repetitions in actual wireless networks.

6.4.2. Real Implementation of Coded Slotted Aloha

- Participants: Cédric Adjih, Vinod Kumar

In 2017, we implemented Coded Slotted Aloha (CSA) as a proof of concept on our FIT IoT-LAB testbed (with 20+ nodes), with 802.15.4 transmissions and using a real SDR software.

This was presented in the seminar of the GT task 2 Future Access Networks of Digicosme. It was also presented as part of the tutorial "IoT in practice" in the ANTS 2017 conference.

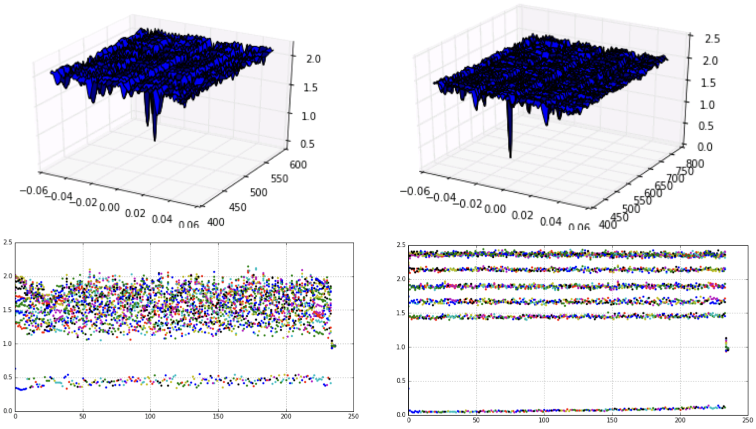


Figure 2. Example of demodulation steps in Coded Slotted Aloha

6.5. Resource and Traffic Management

6.5.1. Utility Optimization Approach to Network Cache Design

Participants: Mostafa Dehghan, Laurent Massoulié, Don Towsley, Daniel Menasche, Y.c. Tay.

In any caching system, the admission and eviction policies determine which contents are added and removed from a cache when a miss occurs. Usually, these policies are devised so as to mitigate staleness and increase the hit probability. Nonetheless, the utility of having a high hit probability can vary across contents. This occurs, for instance, when service level agreements must be met, or if certain contents are more difficult to obtain than others. In this paper, we propose utility-driven caching, where we associate with each content a utility, which is a function of the corresponding content hit probability. We formulate optimization problems where the objectives are to maximize the sum of utilities over all contents. These problems differ according to the stringency of the cache capacity constraint. Our framework enables us to reverse engineer classical replacement policies such as LRU and FIFO, by computing the utility functions that they maximize. We also develop online algorithms that can be used by service providers to implement various caching policies based on arbitrary utility functions.

This work was published and presented at the IEEE Infocom 2016 conference as ”A Utility Optimization Approach to Network Cache Design”.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

1. Participation to Microsoft Research & Inria Joint Centre, which funds two PhD students (Lennart Gulikers and Remi Varloot).
2. Fujitsu has funded further development of RIOT and sponsored the RIOT Summit 2017.
3. Cisco Systems Silicon Valley has sponsored the RIOT Summit 2017.
4. In the framework of the joint research lab between Nokia Bell Labs and Inria, we participate in the ADR (action de recherche) Network Information Theory.

7.2. GranData

- Participants: Aline Carneiro Viana, Guangshuo Chen, Adriano Di Luzio

Since June 2014, we have a collaboration with GranData (http://grandata.com/), Buenos Aires, Argentina on traffic vs mobility modeling of smartphone users. GranData is a small company that integrates first-party and telco partner data to understand key market trends, to predict customer behavior, and to deliver business results. Its products integrates and analyzes diverse data traces (e.g., telco, social media, or mobile data) to generate behavioral insights and deliver targeted mobile marketing. Part of the thesis of Eduardo Mucelli analysis data traffic using telco traces provided by GranDatas. While this collaboration allow us collaborating with machine learning experts, GranData has the opportunity to get our expertise in mobility analysis.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. Equipex FIT

Participants: Cedric Adjih, Emmanuel Baccelli, Alexandre Abadie, Philippe Lubrano, Ichrak Amdouni, Alaeddine Weslati, Vincent Ladeveze.

Partners: Inria (Lille, Sophia-Antipolis, Grenoble), INSA, UPMC, Institut Telecom Paris, Institut Télécom Evry, LSIIT Strasbourg.

Figure 3. FIT IoT-LAB site in Saclay

FIT (Future Internet of Things) aims to develop an experimental facility, a federated and competitive infrastructure with international visibility and a broad panel of customers. It provides this facility with a set of complementary components that enable experimentation on innovative services for academic and industrial users. The project gives french internet stakeholders a means to experiment on mobile wireless communications at the network and application layers thereby accelerating the design of advanced networking technologies for the future internet. FIT was one of 52 winning projects from the first wave of the French Ministry of Higher Education and Research's "Équipements d'Excellence" (Equipex) research grant program, in 2011.

One component of the FIT platform is the sets of IoT-LAB testbeds (see the IoT-LAB web site). These were motivated by the observation that the world is moving towards an “Internet of Things”, in which most communication over networks will be between objects rather than people.

The Infine team is more specifically managing the FIT IoT-LAB site formerly at Rocquencourt, which recently moved to Saclay (on-going re-deployment), and is participating in the deployment of an additional IoT-lab testbed in Berlin (at Freie Universitaet Berlin).

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

8.2.1.1. AGILE (H2020 project)

Participants: Emmanuel Baccelli, Cedric Adjih.

Program: H2020 ICT-30-2015 Topic: Internet of Things and Platforms for Connected Smart Objects

Project acronym: AGILE

Project title: Adoptive Gateways for dIverse muLtiple Environments

Duration: 2015-2017

Coordinator: Emmanuel Baccelli

Other partners: Canonical (UK), Eclipse IoT Foundation (IE), Mobistar (BE), Libelium (ES), Startupbootcamp IoT (SP), CREATE-NET (IT), iMinds (BE), Atos (SP), Rulemotion (UK), Jolocom (DE), Passau University (DE), Sky-Watch (DN), BioAssist (GR), Graz Technical University (AT), Eurotech (IT), IoTango (US).

Abstract:

The AGILE project is a 3-year H2020 project started in January 2016, which will deliver an integrated framework of open source tools and platforms that interoperate for enabling the delivery of adaptive, self-configurable and secure IoT elements (both software and hardware) that can be utilized in a variety of scenarios. Such tools target actors with heterogeneous skills, including entrepreneurs, researchers, and individuals, aiming to enable the realization of IoT applications respecting user privacy and data ownership.

8.2.1.2. ARMOUR (H2020 project)

Participants: Emmanuel Baccelli, Cedric Adjih.

Program: H2020 ICT-12-2015 Topic: Integrating experiments and facilities in FIRE+

Project acronym: ARMOUR

Project title: Large-Scale Experiments of IoT Security Trust

Duration: 2016-2018

Coordinator: Serge Fdida (UPMC)

Other partners: UPMC (France), Synelixis (Greece), SMA (France), UI (Portugal), JRC (Belgium), EGM (France), OdinS (Spain).

Abstract: The ARMOUR project is a 2-year H2020 project started in February 2016. The ARMOUR project is aimed at providing duly tested, benchmarked and certified Security & Trust technological solutions for large-scale IoT using upgraded FIRE large scale IoT/Cloud testbeds properly-equipped for Security & Trust experimentations. To this, ARMOUR will: (1) Enhance two outstanding FIRE testbeds (> 2700 nodes; 500 users) with the ARMOUR experimentation toolbox for enabling large-scale IoT Security & Trust experiments; (2) Deliver six properly experimented, suitably validated and duly benchmarked methods and technologies for enabling Security & Trust in the large-scale IoT; and (3) Define a framework to support the design of Secure & Trusted IoT applications as well as establishing a certification scheme for setting confidence on Security & Trust IoT solutions.

8.2.2. Collaborations with Major European Organizations

8.2.2.1. EU CHIST-ERA MACACO

Participants: Aline Carneiro Viana, Emmanuel Baccelli, Eduardo Mucelli.

Program: EU CHIST-ERA, topic Context- and Content-Adaptive Communication Networks

Project acronym: MACACO

Project title: Mobile context-Adaptive CAching for COntent-centric networking

Duration: 2013-2017 (extended until December 2017)

Coordinator: Aline Carneiro Viana

Other partners: INPT-ENSEEIHT at University of Toulouse, University of Birmingham (UK), SUPSI (Switzerland), CNR (Italy) and Federal University of Minas Gerais (Brazil)

Abstract:

MACACO (Mobile context-Adaptive CAching for COntent-centric networking) is a 3-year CHIST-ERA European Project addressing the topic Context- and Content-Adaptive Communication Networks. Due to delay in funding access and data collection campaign we got an extension until December 2017. It is funded by ANR in France, SNSF in Switzerland, and ESPRC in UK. It focus on data offloading mechanisms that take advantage of context and content information. Our intuition is that if it is possible to extract and forecast the behaviour of mobile network users in the three dimensional space of time, location and interest (i.e. what, when and where users are pulling data from the network), it is possible to derive efficient data offloading protocols. Such protocols would pre-fetch the identified data and cache it at the network edge at an earlier time, preferably when the mobile network is less charged, or offers better quality of service. This project has officially started in November 2013.

8.3. International Initiatives

8.3.1. Inria Associate Teams Not Involved in an Inria International Labs

8.3.1.1. EMBRACE

Title: Leveraging Human Behavior and Uncertainty in 5G Networks to Build Robust Resource Allocation and Services Orchestration Models

International Partners (Institution - Laboratory - Researcher):

UTFPR (Brazil) - Departamento Academico de Informitica (DAINF) Curso de Pos-Graduacao em Engenharia Eletrica e Informatica Industrial (CPGEI) - Anelise Munaretto

UFG (Brazil) - Institute of Computational Mathematics and Scientific / Engineering Computing - Kleber Vieira Cardoso

UFMG (Brazil) - Dpt of Statistics - Antonio A. F. Loureiro

Start year: 2017

See also: https://team.inria.fr/infine/embrace/

EMBRACE propose une architecture novatrice pour gérer des ressources et des services opérationnels hétérogénes. EMBRACE se concentre sur les défis scientifiques liés des ensembles de données collectées dans le monde réel et décrivant le comportement du réseau des utilisateurs. En particulier, EMBRACE exploite la modélisation du comportement humain en termes de mobilité, de demande de contenu, d'intérêts communs et des interactions entre-utisateurs. En construisant des modèles d'allocation les ressources tenant compte de l'utilisateur, EMBRACE a pour objectif de diminuer l'incertitude et mieux cerner les profils humains dans les rï¿œseaux 5G. La communication D2D sera également utilisée comme service opérationnel pour gérer la croissance du trafic mobile en libérant des ressources des réseaux cellulaires, sans augmenter les couts. La nouveauté de l'architecture réside dans les algorithmes concus qui exploiteront les caractérisations tirés de l'analyse du comportement des utilisateurs, l'hétérogénéité du réseau, et de l'incertitude. L'évaluation par simulation et l'émulation sera également l'un des thèmes clés. Enfin, les équipes concernées (Inria Infine, UFMG, UFG, UTFPR) ont un long historique de coopération sur ces thèmes.

8.3.2. Inria International Partners

8.3.2.1. Declared Inria International Partners

1. Renewed IOTPUSH collaboration with Freie Universitaet Berlin around the long-term stay of Emmanuel Baccelli in Berlin, on research topics about the Internet of Things, RIOT and Information-Centric Networking.

8.3.2.2. Informal International Partners

1. On-going collaboration with Freie Universitaet Berlin and Hamburg University of Applied Science around RIOT.
2. Informal collaborations with UIUC and UMass.
3. Informal collaborations with ENSI Tunis and Sesame Tunis.
4. On-going strong collaboration with Sapienza University of Rome, Italy.
5. On-going strong collaboration with CNR Torino, Italy.
6. On-going collaboration with University of Porto, Portugal.
7. On-going collaboration with ENSAE/CNRS, France.
8. On-going collaboration with University of Edinburgh, UK.

8.3.3. Participation in Other International Programs

8.3.3.1. Indo-French project

The Inria teams Infine and Eva are part of the "D2D Communication for LTE Advanced Cellular Network" , a project funded by the Indo-French Centre for the Promotion of Advanced Research (CEFIPRA). With industrial partners, and also with Indian partners, this project is focusing on the evolution of cellular networks towards 5G: this includes exploration of device-to-device (D2D) communication, and more generally IoT communication in a cellular context. Research directions include efficient access for IoT devices (massive numbers of devices with low volume communication); combination of random access protocols/error coding/physical layer ; efficient neighbor discovery, ...

8.3.3.2. STIC AmSud MOTIf 2017

Participant: Aline Carneiro Viana.

Program: STIC AmSud

Project title: Mobile phone sensing of human dynamics in techno-social environment

Duration: 2017-2019

Coordinators: Marton Karsai (ENS/Inria) and Jussara M. Almeida (UFMG) and Alejo Salles (Univ. of Buenos Aires)

Abstract: Information and Communication Technology (ICT) is becoming increasingly social, as demonstrated by the multitude of emerging technologies and technology platforms that facilitate social interactions, taking place as communication via telephone, text message, email, online social networks etc. At the same time, our social activities are increasingly embedded in the ICT environments that enable and enhance our ability to transact, share experiences, and maintain social relationships. One of the best ways to explore these developments is through the mining and analysis of data, which are collected through mobile phones and allow us to investigate how individuals act when embedded in a technology-enabled environment. The MOTIf project builds on the analysis and modeling of geo-localized temporally detailed but fully anonymised mobile phone call networks. These datasets allow us to address the two scientific objectives about spatiotemporal patterns of service usage of anonymised individuals to learn when, where, and what people are doing; and about the fine-grained sociodemographic structure of society and its effect on the the individual social behaviour. In other words our goal in general is to understand how individuals behave in a dynamic techno-social environment.

8.4. International Research Visitors

8.4.1. Visits of International Scientists

Rik Sarkar was Visiting Researcher at Infine for 3 days. He worked with Aline C. Viana and the internship Maria Astefanoaei on predicting new places to visit in human mobility decision.

Julinda Stefa was Visiting Researcher at Infine for 3 months. She worked with Aline C. Viana and the internship Adriano Di Luzio on the inference of human personality from mobile phones datasets.

Ana Aguiar was Visiting Researcher at Infine for 3 days. She worked with Aline C. Viana and the internship Emanuel Lima on data offloading decision via mobile crowdsensing.

8.4.1.1. Internships

Panagiota Katsikouli did an internship of 5 months at Infine working with Aline C. Viana and Marco Fiore on sampling frequency of human mobility.

Maria Astefanoaei did an internship of 5 months at Infine working with Aline C. Viana and Rik Sarkar on predicting new places to visit in human mobility decision.

Adriano Di Luzio did an internship of 4 months at Infine working with Aline C. Viana and Julinda Stefa on the inference of human personality from mobile phones datasets.

Emanuel Lima did an internship of 3 months at Infine working with Aline C. Viana and Ana Aguiar on on data offloading decision via mobile crowdsensing.

Ayat Zaki Hindi did an internship of 6 months at Infine working with Cedric Adjih, Michel Kieffer and C. Weidmann on synchronization strategy in Information-Centric Networks.

8.4.2. Visits to International Teams

8.4.2.1. Research Stays Abroad

Emmanuel Baccelli is Visiting Professor at Freie Universitaet (FU) Berlin, within the context of the formal collaboration IOTPUSH with this university on research topics about the Internet of Things, RIOT and Information-Centric Networking.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

1. **Aline C. Viana** was scientific chair of Algotel 2017.
2. **Emmanuel Baccelli** was chair of the RIOT Summit 2017.

9.1.1.2. Member of the Organizing Committees

1. **Aline C. Viana** is Publicity Chair of: IEEE PICom 2018; IEEE MiSeNet 2018 jointly with IEEE Infocom 2018; DCOSS 2018.

9.1.2. Scientific Events Selection

9.1.2.1. Chair of Conference Program Committees

1. **Aline C. Viana** was scientific chair of Algotel 2017.

9.1.2.2. Member of the Conference Program Committees

1. **Aline C. Viana** is/was TPC member of: Altogel 2018; IFIP NTMS 2018 New Technolog; WCU 2017 jointly with SBRC 2017; IEEE AINA-2017.

9.1.2.3. Reviewer

1. **Cedric Adjih** was reviewer for PEMWN 2017, IEEE PIMRC 2017 Workshop WS-07, WINCOM 2017, WCNC 2018.
2. **Emmanuel Baccelli** was reviewer for ACM ICN 2017 (Poster/Demo), PEMWN 2017.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

1. **Aline C. Viana** is Associate Editor of ACM SIGCOMM Computer Communication Review (ACM CCR), since May 2014; Editorial Board member of Wireless Communications and Mobile Computing Open Access Journal of John Wiley&Sons and Hindawi since 2016.

9.1.3.2. Reviewer - Reviewing Activities

1. **Aline C. Viana** reviewed papers for ACM SIGCOMM CCR Journal, elsevier Pervasive and Mobile Computing Journal and Computer Communication Elsevier Journal.

9.1.3.3. Reviewer

1. **Cedric Adjih** was reviewer for Future Generation Computer Systems, and IEEE Transactions on Communications.
2. **Emmanuel Baccelli** was reviewer for ACM Transactions on Networking (TON), Elsevier Future Generation Computer Systems.

9.1.4. Standardization

1. **Emmanuel Baccelli and Cedric Adjih** have participated at several working groups at IETF (March, July and November 2017)
2. **Emmanuel Baccelli and Cedric Adjih** have participated at the IETF hackathons (July and November 2017)

9.1.5. Invited Talks

1. **Aline C. Viana** was invited to give a talk at (1) University of Coimbra, Portugal in December 2016; (2) University of Porto, Portugal in December 2016; (3) EBSIS Summer School in Timmendorfer Strand, Germany in July 2017; (4) University of Edinburgh, UK in February 2017, (5) Spring School on Networks (SSN) 2017 at ChileCON Conference, Pucon, Chile in October 2017, all on "Toward a more Tactful Networking".
2. **Aline C. Viana** will be the keynote speaker of the CoUrb 2018 Workshop jointly organized with SBRC 2018, Brazil in May 2018.
3. **Laurent Massoulie** gave several invited talks at: Stochastic Networks Conference, UCSD; CIRM workshop on random matrices; Institut Henri Poincarï¿œs Nexus of Information and Computation Theories; LINCS scientific advisory board.
4. **Cedric Adjih** has presented in the June seminar of Digicosme GT task 2 Future Access Networks : “On Doing 5G with Old Low Cost Wireless Sensor Nodes (with RIOT, FIT IoT-LAB and one SDR device)”. This presented work-in-progress experimenting of some 5G-related method(s) on low cost wireless sensor nodes, focusing on massive access (for machine-type communication), and describing one of the most readily experimental family of methods for random access known as "Coded Slotted Aloha", along with its already known features and performance evaluation aspects. Proof-of-concept experimentation was also described starting from the underlying hardware platform, FIT IoT-LAB (a large scale open testbed IoT Lab), and the underlying software platform, RIOT (an Operating System from the Internet of Things). Using additionally a software-defined receiver, it was shown how such 5G-like random access can be experimented on such typical 802.15.4-based devices. An analysis of preliminary results was provided, along with some lessons learned and some consequences on the design space of such methods.
5. **Cedric Adjih** has presented a tutorial at ANTS 2017 in Bhubaneswar, Odisha, India, on the subject of "IoT in Practice": it presented a comparative study of embedded OS?s such as RIOT, Contiki, ARM Mbed, Zephyr as well as implementation specificities of IoT protocol stacks e.g. RPL/6LoWPAN. Insight on how do hardware and software building blocks and transport level and application level protocols fit together was be provided. Addressed practical issues include the analysis of actual protocol performance (e.g. 802.15.4 or LoRa stacks) in experimental settings, and topics such as reasons for adopting open source approaches for IoT solutions. Presentation of experimental proof of concepts concluded the tutorial.

6. **Emmanuel Baccelli** gave an invited talk at CiscoX Symposium 2017, on the topic of “ICN for IoT: Energy Efficiency, Opportunities & Challenges”
7. **Emmanuel Baccelli** gave an invited talk at Bremen University in February, on the topic of “End-to-End Open Source IoT with RIOT”.

9.1.6. Research Administration

1. **Aline C. Viana** is the President of the Scientific Commission at Inria Saclay, responsible for the selection of candidates for the CORDI-S, Post-Doc and Delegation campaigns.
2. **Aline C. Viana** will be the new leader of Infine team due to the fact that Laurent Massoulie had to leave Inria Saclay to join the Inria & Microsoft Lab at Inria Paris.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master : Emmanuel Baccelli, “Operating Systems for Small Connected Devices in the Internet of Things”, 3h cours magistral, Formation PESTO Corps des Mines, Telecom ParisTech, Paris France

Licence: Aline C. Viana, “Toward a more Tactful Networking”, 2h TD, niveau L3, Federal University of Goias (UFG) and Federal University of Minas Gerais, Brazil

Licence: Aline C. Viana, “Opportunistic communication in disconnected mobile Ad Hoc networks”, 6h TD, niveau L3, Telecom SudParis, France

Licence: Cédric Adjih, “Microcontrollers: from embedded systems to the Internet of Things”, 3h TD, Niveau L3, Telecom SudParis, France

Licence: Laurent Massoulie, “Networks: distributed control and emerging phenomena”, 18h TD, Niveau L3, Ecole Polytechnique, France

Master : Aline C. Viana, “Toward a more Tactful Networking”, 2h TD, niveau M2, Federal University of Goias (UFG), Federal University of Minas Gerais, Brazil

Master : Aline C. Viana, “Toward a more Tactful Networking”, 3h TD, niveau M2, Telecom SudParis, France

Master : Aline C. Viana, Master Project evaluation, 4h TD, niveau M2, Telecom SudParis, France

Doctorat : Aline C. Viana, “Toward a more Tactful Networking”, 2h TD, Federal University of Goias (UFG) and Federal University of Minas Gerais, Brazil

9.2.2. Supervision

PhD: Guangshuo Chen, “Human habits investigation: From mobility reconstruction to mobile traffic prediction”, Ecole Polytechnique, March 2018, Aline C. Viana.

PhD in progress: (1) Rafael Costa, “Infering friends in the crowd in Device-to-Device communication”, June 2016, Aline C. Viana and Leobino Sampaio (co-tutele); (2) Roni Shigetta, “Distributed channel allocation strategy based on human behavior”, January 2014, Aline C. Viana and Mauro Fonseca (co-tutele).

PhD in progress: Loic Dauphin, “IoT Robotics”, June 2016, Emmanuel Baccelli and Cédric Adjih co-supervision.

9.2.3. Juries

1. Aline Viana will be on the PhD jury as reviewer of: Florent Coriat, “Geolocalisation en situation de crise”, UPMC ? Sorbonne Universitï¿œs, France, may 2018.

2. Aline Viana was on the PhD jury as examiner of: (1) S. Eddine Belouanas, “Dessemination de contenus populaires et tolerants au delai dans les reseaux cellulaires”, UPMC ? Sorbonne Universitï¿½s, France, September 2017; (2) P. Salgueiro Santos,“Wireless Protocols and Channel Estimation for Data Gathering with Mobile Nodes”, Univ. of Porto, Portugal, May 2017; (3) Amal Ellouzem, "Mobile Applications Offloading in Mobile Cloud Environment", Telecom ParisTech, France, March 2017.
3. Aline C. Viana was on the PhD jury as invited member of Tanel Razafimandimby "Toward Internet of Heterogeneous Things: Wireless communication maintenance and efficient data sharing among devices", University Lille 1, Inria Lille, France, October 2017.
4. Aline C. Viana was on the PhD Qualification jury as examiner of T. Duc Ha, "Allocation de ressources et association utilisateur/cellule optimisees pour les reseaux C-RAN" (Univ. Paris Sud, 2017);
5. Aline C. Viana was on the Master jury of Bruno Farias Fausto "Um mecanismo de deteccao e controle de congestionamento usando redes ad hoc veiculares sem infraestrutura", Federal Univ. of Rio de Janeiro State, Brazil, October 2017.
6. Laurent Massoulié was on the PhD jury as reviewer for the PhD theses of Anna Benhamou and Alaa Saade, and he presided the PhD thesis committee of Kevin Scaman.

9.3. Popularization

1. Aline C. Viana gave a scientific talk on “Le smartphone, votre 6e sens ?” and “The research career”, December 2017, for college students visiting Inria Saclay
2. Cédric Adjih, Emmanuel Baccelli, Alexandre Abadie, Raul Fuentes participated to the "Inria-Industry Meetings" on 17 and 18 October 2017, with two demonstrations:
   – RIOT, the friendly OS for the Internet of Things - it has shown the various functionalities from RIOT on hardware of the type sensor/actuator to demo a control loop via the cloud. RIOT was used to generate IoT data locally (via sensors), sending them to the cloud where this data is analyzed. This analysis will result in a physical action (via actuators). We highlighted two use-cases: building automation, and low-cost robotics
   – End-to-end IoT with FIT IoT-LAB - this presented the use of FIT IoT-LAB for experimenting with IoT; in particular the end-to-end Over-the-air updates of RIOT.

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] L. GULIKERS, M. LELARGE, L. MASSOULIÉ. *A spectral method for community detection in moderately sparse degree-corrected stochastic block models*, in "Advances in Applied Probability", June 2017, vol. 49, n^{o} 03, p. 686 - 721 [*DOI :* 10.1017/APR.2017.18], https://hal.archives-ouvertes.fr/hal-01622731.

[2] E. MUCELLI REZENDE OLIVEIRA, A. CARNEIRO VIANA, K. P. NAVEEN, C. SARRAUTE. *Mobile Data Traffic Modeling: Revealing Temporal Facets*, in "Computer Networks", January 2017, vol. 112, p. 176-193, https://hal.inria.fr/hal-01453379.

Invited Conferences

[3] G. CHEN, A. CARNEIRO VIANA, C. SARRAUTE. *Towards an Adaptive Completion of Sparse Call Detail Records for Mobility Analysis*, in "Workshop on Data Analytics for Mobile Networking", Kona, United States, March 2017, https://hal.inria.fr/hal-01448822.

International Conferences with Proceedings

[4] G. Chen, S. Hoteit, A. Carneiro Viana, M. Fiore, C. Sarraute.*The Spatiotemporal Interplay of Regularity and Randomness in Cellular Data Traffic*, in "LCN 2017 - The 42nd IEEE Conference on Local Computer Networks", Singapore, Singapore, October 2017, https://hal.inria.fr/hal-01646359.

[5] L. Dauphin, H. Petersen, C. Adjih, E. Baccelli.*Low-Cost Robots in the Internet of Things: Hardware, Software & Communication Aspects*, in "NextMote Workshop - Next Generation Platforms for the Cyber-Physical Internet", Uppsala, Sweden, February 2017, The workshop is hosted by the International Conference on Embedded Wireless Systems and Networks (EWSN), https://hal.inria.fr/hal-01666441.

[6] O. Hahm, E. Baccelli, T. C. Schmidt, M. Wählisch, C. Adjih, L. Massoulié.*Low-power Internet of Things with NDN & Cooperative Caching*, in "ACM ICN 2017 - 4th ACM Conference on Information-Centric Networking", Berlin, Germany, September 2017, https://hal.inria.fr/hal-01666434.

[7] P. Katsikouli, A. Carneiro, M. Fiore, A. Tarable.*On the Sampling Frequency of Human Mobility*, in "IEEE GLOBECOM 2017", Singapore, Singapore, December 2017, p. 1-6, https://hal.inria.fr/hal-01651856.

[8] E. E. Khaleghi, C. Adjih, A. Alloum, P. Muhlethaler.*Near-Far Effect on Coded Slotted ALOHA*, in "PIMRC 2017 - IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications - Workshop WS-07 on "The Internet of Things (IoT), the Road Ahead: Applications, Challenges, and Solutions"", Montreal, Canada, October 2017, https://hal.inria.fr/hal-01675805.

[9] P. Kietzmann, C. Gündogan, T. C. Schmidt, O. Hahm, M. Wählisch.*The Need for a Name to MAC Address Mapping in NDN: Towards Quantifying the Resource Gain*, in "ACM ICN 2017 - 4th ACM Conference on Information-Centric Networking", Berlin, Germany, September 2017, https://hal.inria.fr/hal-01666601.

[10] R. F. Shigueta, M. Fonseca, A. Carneiro Viana.*A Mobility-Aware Channel Allocation Strategy for Clustered Ad hoc Network*, in "VTC2017-Spring - 2017 IEEE 85th Vehicular Technology Conference", Sydney, Australia, June 2017, https://hal.inria.fr/hal-01676499.

Conferences without Proceedings

[11] L. Gulikers, M. Lelarge, L. Massoulié.*Non-Backtracking Spectrum of Degree-Corrected Stochastic Block Models*, in "ITCS 2017 - 8th Innovations in Theoretical Computer Science", Berkeley, United States, January 2017, https://arxiv.org/abs/1609.02487 , https://hal.archives-ouvertes.fr/hal-01622719.

[12] V. Shah, L. Gulikers, L. Massoulié, M. Vojnovic.*Adaptive Matching for Expert Systems with Uncertain Task Types*, in "Allerton 2017 - 55th Annual Allerton Conference on Communication, Control, and Computing", Monticello, IL, United States, October 2017, https://arxiv.org/abs/1703.00674 , https://hal.archives-ouvertes.fr/hal-01622739.

Research Reports

[13] G. Chen.*Spatiotemporal Individual Mobile Data Traffic Prediction*, Inria Saclay - Ile-de-France, January 2018, n° RT-0497, https://hal.inria.fr/hal-01675573.

[14] G. CHEN, S. HOTEIT, A. CARNEIRO VIANA, M. FIORE, C. SARRAUTE.*Enriching Sparse Mobility Information in Call Detail Records*, Inria Saclay - Ile-de-France, November 2017, n^{o} RT-0496, https://hal.inria.fr/hal-01646608.

[15] G. CHEN, S. HOTEIT, A. CARNEIRO VIANA, M. FIORE, C. SARRAUTE.*Spatio-Temporal Predictability of Cellular Data Traffic*, Inria Saclay - Ile-de-France, January 2017, n^{o} RT-0483, 17, https://hal.inria.fr/hal-01393361.

[16] G. CHEN, S. HOTEIT, A. CARNEIRO VIANA, M. FIORE, C. SARRAUTE.*Individual Trajectory Reconstruction from Mobile Network Data*, Inria Saclay - Ile-de-France, January 2018, n^{o} RT-0495, p. 1-23, https://hal.inria.fr/hal-01675570.

[17] P. KATSIKOULI, A. C. CARNEIRO VIANA, M. FIORE, A. TARABLE.*On The Sampling Frequency of Human Mobility*, Inria, February 2017, n^{o} RT-0487, 19, https://hal.inria.fr/hal-01470393.

[18] R. LIMA COSTA, A. CARNEIRO VIANA, A. ZIVIANI, L. NASCIMENTO SAMPAIO.*Tactful Networking: Humans in the communication loop* , Inria Saclay - Ile-de-France, January 2018, https://hal.inria.fr/hal-01675445.

Other Publications

[19] B. ADAMSON, C. ADJIH, J. BILBAO, V. FIROIU, F. FITZEK, G. A. M. SAMAH, E. LOCHIN, A. MASUCCI, M.-J. MONTPETIT, M. V. PEDERSEN, G. PERALTA, V. ROCA, S. PARESH, S. SIVAKUMAR.*Network Coding Taxonomy*, July 2017, Internet Research Task Force - Working document of the Network Coding Research Group (NWCRG), draft-irtf-nwcrg-network-coding-taxonomy-05 (work in progress), https://datatracker.ietf.org/doc/draft-irtf-nwcrg-network-coding-taxonomy/, https://hal.inria.fr/hal-00998506.

[20] C. ADJIH, M. KIEFFER, C. GRECO.*Network Coding with Random Packet-Index Assignment for Mobile Crowdsensing*, June 2017, working paper or preprint, https://hal-centralesupelec.archives-ouvertes.fr/hal-01538115.

[21] L. DAUPHIN, E. BACCELLI, C. ADJIH, H. PETERSEN.*NDN-based IoT Robotics*, September 2017, ACM ICN'17 - 4th ACM Conference on Information-Centric Networking, Poster, https://hal.inria.fr/hal-01666539.

[22] V. ROCA, J. DETCHART, C. ADJIH, M. V. PEDERSEN, I. SWETT.*Generic Application Programming Interface (API) for Window-Based Codes*, October 2017, Internet Research Task Force - Working document of the Network Coding Research Group (NWCRG), draft-roca-nwcrg-generic-fec-api-00 (work in progress), https://datatracker.ietf.org/doc/draft-roca-nwcrg-generic-fec-api/, https://hal.inria.fr/hal-01630138.

[23] A. ZAKI HINDI, M. KIEFFER, C. ADJIH, C. WEIDMANN.*NDN Synchronization: iRoundSync, an Improved RoundSync*, September 2017, ICN 2017: 4th ACM Conference on Information-Centric Networking, Poster [*DOI :* 10.1145/3125719.3132108], https://hal.inria.fr/hal-01675930.

Project-Team LIFEWARE

Computational systems biology and optimization

RESEARCH CENTER
Saclay - Île-de-France

THEME
Computational Biology

Table of contents

Project-Team LIFEWARE

Creation of the Team: 2014 January 01, updated into Project-Team: 2015 April 01

Keywords:

Computer Science and Digital Science:

A2.1.1. - Semantics of programming languages
A2.1.5. - Constraint programming
A2.1.10. - Domain-specific languages
A2.2.1. - Static analysis
A2.3.2. - Cyber-physical systems
A2.4. - Verification, reliability, certification
A2.4.1. - Analysis
A2.4.2. - Model-checking
A2.4.3. - Proofs
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.2. - Stochastic Modeling (SPDE, SDE)
A6.1.3. - Discrete Modeling (multi-agent, people centered)
A6.1.4. - Multiscale modeling
A6.2.4. - Statistical methods
A6.2.6. - Optimization
A6.3.1. - Inverse problems
A6.3.4. - Model reduction
A7.2. - Logic in Computer Science
A8.1. - Discrete mathematics, combinatorics
A8.2. - Optimization
A8.7. - Graph theory
A9.7. - AI algorithmics

Other Research Topics and Application Domains:

B1. - Life sciences
B1.1.2. - Molecular biology
B1.1.3. - Cellular biology
B1.1.9. - Bioinformatics
B1.1.10. - Mathematical biology
B1.1.11. - Systems biology
B1.1.12. - Synthetic biology
B1.1.14. - Microbiology
B2. - Health
B2.2.3. - Cancer
B2.4.1. - Pharmaco kinetics and dynamics
B2.4.2. - Drug resistance
B9. - Society and Knowledge

1. Personnel

Research Scientists

François Fages [Team leader, Inria, Senior Researcher, HDR]
Grégory Batt [Inria, Senior Researcher, HDR]
Jakob Ruess [Inria, Researcher]
Sylvain Soliman [Inria, Researcher, HDR]

Faculty Member

Philippe Dague [Univ Paris-Sud, Professor, from Feb 2017 until Aug 2017]

External Collaborator

Denis Thieffry [Ecole Normale Supérieure Paris, HDR]

Technical Staff

David Coudrin [Inria, from Nov 2017]
Steven Fletcher [Inria, from Apr 2017]

PhD Students

Chetan Aditya [Inria, from Nov 2017]
Virgile Andreani [Ecole Polytechnique]
Jean Baptiste Caron [Inria, until Feb 2017]
Sebastian Ramon Sosa Carrillo [Inria, from Aug 2017]

Post-Doctoral Fellows

Jean-Baptiste Lugagne [Inria, until Jan 2017, then MSC lab (CNRS/Paris7) until Aug 2017]
Chiara Fracassi [Inria, until Aug 2017]

Administrative Assistant

Corinne Petitot [Inria]

2. Overall Objectives

2.1. Overall Objectives

This project aims at developing formal methods and experimental settings for **understanding the cell machinery** and establishing formal paradigms in cell biology. It is based on the vision of **cells as machines**, **biochemical reaction systems as programs**, and on the use of concepts and tools from computer science to master the complexity of cell processes. While for the biologist, as well as for the mathematician, the size of the biological networks and the number of elementary interactions constitute a complexity barrier, for the computer scientist the difficulty is not that much in the size of the networks than in the unconventional nature of biochemical computation. Unlike most programs, biochemical reaction systems involve transitions that are stochastic rather than deterministic, continuous-time rather than discrete-time, poorly localized in compartments instead of well-structured in modules, and created by evolution instead of by rational design. It is our belief however that some form of modularity is required by an evolutionary system to survive, and that the elucidation of these modules in biochemical computation is now a key to apply engineering methods in cell biology on a large scale.

Concretely, we keep developing a theory of biochemical computation with a prototype implementation in the Biochemical Abstract Machine BIOCHAM, a modeling and analysis platform for Systems Biology. The reaction rule-based language used in this system allows us to reason about biochemical reaction networks at different levels of abstraction, in either the **stochastic, differential, discrete, logical or hybrid semantics** of the reactions. This makes it possible to apply a variety of **static analysis** methods, before going to simulations and **dynamical analyses**, for which we use **temporal logics** as a specification language to formalize biological behaviours with imprecise data, validate models w.r.t. observations, constrain model building, and calibrate models in high dimension by optimization methods

A **tight integration between dry lab and wet lab** efforts is also essential for the success of the project. In collaboration with biologists, we investigate concrete biological questions and develop computational models fitted to quantitative data which allow us to make quantitative predictions. In collaboration with Pascal Hersen, MSC lab, we contribute to the development of an experimental platform for the closed-loop control of intracellular processes. This platform combines hardware (microfluidic device and microscope), software (cell tracking and model-based predictive control algorithms) and genetically modified living cells. It is used to investigate the possibilities to externalize the control of intracellular processes for systems and synthetic biology applications, and perform accurate observations, modifications and real-time control at both single cell and cell population levels. We are affiliated with the Doctorate Schools "Frontières du vivant (FdV)" of University Sorbonne Paris Cité and "Sciences et technologies de l'information et de la communication (STIC)" of University Paris-Saclay.

This project addresses fundamental research issues in computer science on the interplay between **structure and dynamics** in large interaction networks, and on the mixing of **continuous and discrete computation**. Many static analysis problems of biological networks are NP-hard. The recourse to constraint logic programming (CLP) to model and solve them, is our secret weapon, which probably explains our capability to experiment ideas in computational systems biology in very short time, by implementing them in CLP, integrating them as new components in our modeling platform BIOCHAM, and evaluating them directly on a large scale in systems biology model repositories such as BIOMODELS.NET.

3. Research Program

3.1. Computational Systems Biology

Bridging the gap between the complexity of biological systems and our capacity to model and **quantitatively predict system behaviors** is a central challenge in systems biology. We believe that a deeper understanding of the concept and theory of biochemical computation is necessary to tackle that challenge. Progress in the theory is necessary for scaling, and enabling the application of static analysis, module identification and decomposition, model reductions, parameter search, and model inference methods to large biochemical reaction systems. A measure of success on this route will be the production of better computational modeling tools for elucidating the complex dynamics of natural biological processes, designing synthetic biological circuits and biosensors, developing novel therapy strategies, and optimizing patient-tailored therapeutics.

Progress on the **coupling of models to data** is also necessary. Our approach based on quantitative temporal logics provides a powerful framework for formalizing experimental observations and using them as formal specification in model building. Key to success is a tight integration between *in vivo* and *in silico* work, and on the mixing of dry and wet experiments, enabled by novel biotechnologies. In particular, the use of microfluidic devices makes it possible to measure behaviors at both single-cell and cell population levels *in vivo*, provided innovative modeling, analysis and control methods are deployed *in silico*.

In synthetic biology, while the construction of simple intracellular circuits has shown feasible, the design of larger, **multicellular systems** is a major open issue. In engineered tissues for example, the behavior results from the subtle interplay between intracellular processes (signal transduction, gene expression) and intercellular processes (contact inhibition, gradient of diffusible molecule), and the question is how should cells be genetically modified such that the desired behavior robustly emerges from cell interactions.

3.2. Chemical Reaction Network (CRN) Theory

Feinberg's chemical reaction network theory and Thomas's influence network analyses provide sufficient and/or necessary structural conditions for the existence of multiple steady states and oscillations in regulatory networks. Those conditions can be verified by static analyzers without knowing kinetic parameter values nor making any simulation. In this domain, most of our work consists in analyzing the interplay between the **structure** (Petri net properties, influence graph, subgraph epimorphisms) and the **dynamics** (Boolean, CTMC, ODE, time scale separations) of biochemical reaction systems. In particular, our study of influence graphs of reaction systems, our generalization of Thomas' conditions of multi-stationarity and Soulé's proof to reaction systems [0], the inference of reaction systems from ODEs [0], the computation of structural invariants by constraint programming techniques, and the analysis of model reductions by subgraph epimorphisms now provide solid ground for developing static analyzers, using them on a large scale in systems biology, and elucidating modules.

3.3. Logical Paradigm for Systems Biology

Our group was among the first ones in 2002 to apply **model-checking** methods to systems biology in order to reason on large molecular interaction networks, such as Kohn's map of the mammalian cell cycle (800 reactions over 500 molecules) [0]. The logical paradigm for systems biology that we have subsequently developed for quantitative models can be summarized by the following identifications :

$$\text{biological model = transition system } K$$

$$\text{dynamical behavior specification = temporal logic formula } \phi$$

$$\text{model validation = model-checking } \; K,\; s \models ?\; \phi$$

$$\text{model reduction = sub-model-checking, } K' \subset K \text{ s.t. } K'?,\; s \models \phi$$

$$\text{model prediction = formula enumeration, } \phi \text{ s.t. } K,\; s \models \phi?$$

$$\text{static experiment design = symbolic model-checking, state } s \text{ s.t. } K,\; s? \models \phi$$

$$\text{model synthesis = constraint solving } \; K?,\; s \models \phi$$

$$\text{dynamic experiment design = constraint solving } \; K?,\; s? \models \phi$$

In particular, the definition of a continuous satisfaction degree for **first-order temporal logic** formulae with constraints over the reals, was the key to generalize this approach to quantitative models, opening up the field of model-checking to model optimization [0] This line of research continues with the development of temporal logic patterns with efficient constraint solvers and their generalization to handle stochastic effects.

3.4. Modeling of Phenotypic Heterogeneity in Cellular Processes

Since nearly two decades, a significant interest has grown for getting a quantitative understanding of the functioning of biological systems at the cellular level. Given their complexity, proposing a model accounting for the observed cell responses, or better, predicting novel behaviors, is now regarded as an essential step to validate a proposed mechanism in systems biology. Moreover, the constant improvement of stimulation and observation tools creates a strong push for the development of methods that provide predictions that are increasingly precise (single cell precision) and robust (complex stimulation profiles).

It is now fully apparent that cells do not respond identically to a same stimulation, even when they are all genetically-identical. This phenotypic heterogeneity plays a significant role in a number of problems ranging from cell resistance to anticancer drug treatments to stress adaptation and bet hedging.

Dedicated modeling frameworks, notably **stochastic** modeling frameworks, such as chemical master equations, and **statistic** modeling frameworks, such as ensemble models, are then needed to capture biological variability.

[0] Sylvain Soliman. A stronger necessary condition for the multistationarity of chemical reaction networks. Bulletin of Mathematical Biology, 75(11):2289–2303, 2013.

[0] François Fages, Steven Gay, Sylvain Soliman. Inferring reaction systems from ordinary differential equations. Journal of Theoretical Computer Science (TCS), Elsevier, 2015, 599, pp.64–78.

[0] N. Chabrier-Rivier, M. Chiaverini, V. Danos, F. Fages, V. Schächter. Modeling and querying biochemical interaction networks. Theoretical Computer Science, 325(1):25–44, 2004.

[0] On a continuous degree of satisfaction of temporal logic formulae with applications to systems biology A. Rizk, G. Batt, F. Fages, S. Soliman International Conference on Computational Methods in Systems Biology, 251-268

Appropriate mathematical and computational should then be employed for the analysis of these models and their calibration to experimental data. One can notably mention **global optimization** tools to search for appropriate parameters within large spaces, **moment closure** approaches to efficiently approximate stochastic models [0], and (stochastic approximations of) the **expectation maximization** algorithm for the identification of mixed-effects models [0].

3.5. External Control of Cell Processes

External control has been employed since many years to regulate culture growth and other physiological properties. Recently, taking inspiration from developments in synthetic biology, closed loop control has been applied to the regulation of intracellular processes. Such approaches offer unprecedented opportunities to investigate how a cell process dynamical information by maintaining it around specific operating points or driving it out of its standard operating conditions. They can also be used to complement and help the development of synthetic biology through the creation of hybrid systems resulting from the interconnection of in vivo and in silico computing devices.

In collaboration with Pascal Hersen (CNRS MSC lab), we developed a platform for gene expression control that enables to control protein concentrations in yeast cells. This platform integrates microfluidic devices enabling long-term observation and rapid change of the cells environment, microscopy for single cell measurements, and software for real-time signal quantification and model based control. We demonstrated in 2012 that this platform enables controlling the level of a fluorescent protein in cells with unprecedented accuracy and for many cell generations [0].

More recently, motivated by an analogy with a benchmark control problem, the stabilization of an inverted pendulum, we investigated the possibility to balance a genetic toggle switch in the vicinity of its unstable equilibrium configuration. We searched for solutions to balance an individual cell and even an entire population of heterogeneous cells, each harboring a toggle switch [2].

Independently, in collaboration with colleagues from IST Austria, we investigated the problem of controlling cells, one at a time, by constructing an integrated optogenetic platform. It enables experiments that bridge individual and population behaviors. We demonstrated: (i) population structuring by independent closed-loop control of gene expression in many individual cells, (ii) cell–cell variation control during antibiotic perturbation, (iii) hybrid bio-digital circuits in single cells, and freely specifiable digital communication between individual bacteria [1].

3.6. Synthesis of Continuous CRNs

The continuous nature of many protein interactions leads us to consider models of analog computation, and in particular, the recent results in the theory of analog computability and complexity obtained by Amaury Pouly [0] and Olivier Bournez, establish fundamental links with digital computation. In [10], we derive from these results a Turing completeness result for elementary reaction systems (without polymerization) under the differential semantics. The proof of this result shows how mathematical functions described by Ordinary Differential Equations, namely by Polynomial Initial Value Problems (PIVP), can be compiled into elementary biochemical reactions, furthermore with a notion of analog computation complexity defined as the length of the trajectory to reach a given precision on the result. This opens a whole research avenue to analyze natural

[0] Moment-based inference predicts bimodality in transient gene expression, C. Zechner C, J. Ruess, P. Krenn, S. Pelet, M. Peter, J. Lygeros, and H. Koeppl, Proceedings of the National Academy of Sciences USA, 9(5):109(21):8340-5, 2012

[0] What population reveals about individual cell identity: estimation of single-cell models of gene expression in yeast, A. Llamosi, A.M. Gonzalez-Vargas, C. Versari, E. Cinquemani, G. Ferrari-Trecate, P. Hersen, and G. Batt, PLoS Computational Biology, 9(5): e1003056, 2015

[0] Jannis Uhlendorf, Agnés Miermont, Thierry Delaveau, Gilles Charvin, François Fages, Samuel Bottani, Grégory Batt, Pascal Hersen. Long-term model predictive control of gene expression at the population and single-cell levels. Proceedings of the National Academy of Sciences USA, 109(35):14271–14276, 2012.

[0] Amaury Pouly, “Continuous models of computation: from computability to complexity”, PhD Thesis, Ecole Polytechnique, Nov. 2015.

circuits in Systems Biology, transform behavioural specifications into biochemical reactions for Synthetic Biology, and compare artificial circuits with natural circuits acquired through evolution, from the novel point of view of analog computation complexity.

3.7. Constraint Solving and Optimization

Constraint solving and optimization methods are important in our research. On the one hand, static analysis of biochemical reaction networks involves solving hard combinatorial optimization problems, for which **constraint programming** techniques have shown particularly successful, often beating dedicated algorithms and allowing to solve large instances from model repositories. On the other hand, parameter search and model calibration problems involve similarly solving hard continuous optimization problems, for which **evolutionary algorithms** such as the covariance matrix evolution strategy (CMA-ES) [0] has shown to provide best results in our context, for up to 100 parameters, for building challenging quantitative models, gaining model-based insights, revisiting admitted assumptions, and contributing to biological knowledge [0].

4. Application Domains

4.1. Preamble

Our collaborative work on biological applications is expected to serve as a basis for groundbreaking advances in cell functioning understanding, cell monitoring and control, and novel therapy design and optimization. Our collaborations with biologists are focused on **concrete biological questions**, and on the building of predictive models of biological systems to answer them. However, one important application of our research is the development of a **modeling software** for computational systems biology.

4.2. Modeling software for systems biology

Since 2002, we develop an open-source software environment for modeling and analyzing biochemical reaction systems. This software, called the Biochemical Abstract Machine (BIOCHAM), is compatible with SBML for importing and exporting models from repositories such as BioModels. It can perform a variety of static analyses, specify behaviors in Boolean or quantitative temporal logics, search parameter values satisfying temporal constraints, and make various simulations. While the primary reason of this development effort is to be able to **implement our ideas and experiment them quickly on a large scale**, BIOCHAM is used by other groups either for building models, for comparing techniques, or for teaching (see statistics in software section). BIOCHAM-WEB is a web application which makes it possible to use BIOCHAM without any installation. We plan to continue developing BIOCHAM for these different purposes and improve the software quality.

4.3. Couplings between the cell cycle and the circadian clock

Recent advances in cancer chronotherapy techniques support the evidence that there exist important links between the cell cycle and the circadian clock genes. One purpose for modeling these links is to better understand how to efficiently target malignant cells depending on the phase of the day and patient characterictics. These questions are at the heart of our collaboration with Franck Delaunay (CNRS Nice) and Francis Lévi (Univ. Warwick, GB, formerly INSERM Hopital Paul Brousse, Villejuif) and of our participation in the ANR HYCLOCK project and in the submitted EU H2020 C2SyM proposal, following the former EU EraNet Sysbio

[0] N. Hansen, A. Ostermeier (2001). Completely derandomized self-adaptation in evolution strategies. Evolutionary Computation, 9(2) pp. 159–195.

[0] Domitille Heitzler, Guillaume Durand, Nathalie Gallay, Aurélien Rizk, Seungkirl Ahn, Jihee Kim, Jonathan D. Violin, Laurence Dupuy, Christophe Gauthier, Vincent Piketty, Pascale Crépieux, Anne Poupon, Frédérique Clément, François Fages, Robert J. Lefkowitz, Eric Reiter. Competing G protein-coupled receptor kinases balance G protein and β-arrestin signaling. Molecular Systems Biology, 8(590), 2012.

C5SYS and FP6 TEMPO projects. In the past, we developed a coupled model of the Cell Cycle, Circadian Clock, DNA Repair System, Irinotecan Metabolism and Exposure Control under Temporal Logic Constraints [0]. We now focus on the bidirectional coupling between the cell cycle and the circadian clock and expect to gain fundamental insights on this complex coupling from computational modeling and single-cell experiments.

4.4. Biosensor design and implementation in non-living protocells

In collaboration with Franck Molina (CNRS, Sys2Diag, Montpellier) and Jie-Hong Jiang (NTU, Taiwan) we ambition to apply our techniques to the design and implementation of biosensors in non-living vesicles for medical applications. Our approach is based on purely protein computation and on our ability to compile controllers and programs in biochemical reactions. The realization will be prototyped using a microfluidic device at CNRS Sys2Diag which will allow us to precisely control the size of the vesicles and the concentrations of the injected proteins. It is worth noting that the choice of non-living chassis, in contrast to living cells in synthetic biology, is particularly appealing for security considerations and compliance to forthcoming EU regulation.

5. Highlights of the Year

5.1. Highlights of the Year

- **Virtual Reality for Bacteria**

  Individual bacteria have been interfaced with a computer to build hybrid bio-digital circuits. Study published in Nature Communications [1].
- **Dynamical stabilization: real-time control allows maintaining cells in unstable configurations.**

  Using real-time control or periodic forcing one can drive cells towards a region of instability and dynamically maintain them there. Study published in Nature Communications [2].
- **Strong Turing Completeness of Continuous CRNs** solving a long standing open problem in CRN theory [8].

5.1.1. Awards

BEST PAPERS AWARDS :
[8] **CMSB 2017 - 15th International Conference on Computational Methods in Systems Biology**. F. FAGES, G. LE GULUDEC, O. BOURNEZ, A. POULY.

6. New Software and Platforms

6.1. BIOCHAM

The Biochemical Abstract Machine
KEYWORDS: Systems Biology - Bioinformatics
FUNCTIONAL DESCRIPTION: The Biochemical Abstract Machine (BIOCHAM) is a software environment for modeling, analyzing and synthesizing biochemical reaction networks (CRNs) with respect to a formal specification of the observed or desired behavior of a biochemical system. BIOCHAM is compatible with the Systems Biology Markup Language (SBML) and contains some unique features about formal specifications in quantitative temporal logic, sensitivity and robustness analyses and parameter search in high dimension w.r.t. behavioral specifications, static analyses, and synthesis of CRNs.

[0] Elisabetta De Maria, François Fages, Aurélien Rizk, Sylvain Soliman. Design, Optimization, and Predictions of a Coupled Model of the Cell Cycle, Circadian Clock, DNA Repair System, Irinotecan Metabolism and Exposure Control under Temporal Logic Constraints. Theoretical Computer Science, 412(21):2108 2127, 2011.

RELEASE FUNCTIONAL DESCRIPTION: influence networks with forces – PAC learning of influence networks from time series data – synthesis of continuous reaction networks for mathematical functions defined by polynomial differential equations – complete modular rewriting of Biocham in SWI-Prolog

- Participants: François Fages, David Coudrin, Sylvain Soliman and Thierry Martinez
- Contact: François Fages
- URL: http://lifeware.inria.fr/biocham/

7. New Results

7.1. Strong Turing completeness of continuous CRNs

Participants: François Fages, Guillaume Le Guludec (former Member), Sylvain Soliman.

When seeking to understand how computation is carried out in the cell to maintain itself in its environment, process signals and make decisions, the continuous nature of protein interaction processes forces us to consider also analog computation models and mixed analog-digital computation programs. However, recent results in the theory of analog computability and complexity obtained by Pouly and Bournez, establish fundamental links between analog and digital computing. In [8] and [10], we derive from these results the strong (uniform computability) Turing completeness of chemical reaction networks over a finite set of molecular species under the differential semantics, solving a long standing open problem. Furthermore we derive from the proof a compiler of mathematical functions into elementary chemical reactions. We illustrate the reaction code generated by our compiler on trigonometric functions, and on various sigmoid functions which can serve as markers of presence or absence for implementing program control instructions in the cell and imperative programs. This makes it possible to start comparing our compiler-generated circuits to the natural circuit of the MAPK signaling network, which plays the role of an analog-digital converter in the cell with a Hill type sigmoid input/output functions.

7.2. Influence networks compared with reaction networks

Participants: François Fages, Thierry Martinez (former Member), David Rosenblueth (former Member), Sylvain Soliman, Denis Thieffry.

Biochemical reaction networks are one of the most widely used formalism in systems biology to describe the molecular mechanisms of high-level cell processes. However modellers also reason with influence diagrams to represent the positive and negative influences between molecular species and may find an influence network useful in the process of building a reaction network. In [11], we introduce a formalism of influence networks with forces, and equip it with a hierarchy of Boolean, Petri net, stochastic and differential semantics, similarly to reaction networks with rates. We show that the expressive power of influence networks is the same as that of reaction networks under the differential semantics, but weaker under the discrete semantics. Furthermore, the hierarchy of semantics leads us to consider a (positive) Boolean semantics without test for absence, that we compare with the (negative) Boolean semantics with test for absence of gene regulatory networks a` la Thomas. We study the monotonicity properties of the positive semantics and derive from them an algorithm to compute attractors in both the positive and negative Boolean semantics. We illustrate our results on models of the literature about the p53/Mdm2 DNA damage repair system, the circadian clock, and the influence of MAPK signaling on cell-fate decision in urinary bladder cancer.

7.3. Machine learning influence networks from data

Participants: Arthur Carcano, François Fages, Jérémy Grignard, Sylvain Soliman.

Automating the process of model building from experimental data is a very desirable goal to palliate the lack of modellers for many applications. However, despite the spectacular progress of machine learning techniques in data analytics, classification, clustering and prediction making, learning dynamical models from data time-series is still challenging. In [7], we investigate the use of the Probably Approximately Correct (PAC) learning framework of Leslie Valiant as a method for the automated discovery of influence models of biochemical processes from Boolean and stochastic traces. We show that Thomas' Boolean influence systems can be naturally represented by k-CNF formulae, and learned from time-series data with a number of Boolean activation samples per species quasi-linear in the precision of the learned model, and that positive Boolean influence systems can be represented by monotone DNF formulae and learned actively with both activation samples and oracle calls. We consider Boolean traces and Boolean abstractions of stochastic simulation traces, and show the space-time tradeoff there is between the diversity of initial states and the length of the time horizon, together with its impact on the error bounds provided by the PAC learning algorithms. We evaluate the performance of this approach on a model of T-lymphocyte differentiation, with and without prior knowledge, and discuss its merits as well as its limitations with respect to realistic experiments.

7.4. Shaping bacterial population behavior through computer-interfaced control of individual cells

Participant: Jakob Ruess.

Bacteria in groups vary individually, and interact with other bacteria and the environment to produce population-level patterns of gene expression. Investigating such behavior in detail requires measuring and controlling populations at the single-cell level alongside precisely specified interactions and environmental characteristics. In [1], we present an automated, programmable platform that combines image-based gene expression and growth measurements with on-line optogenetic expression control for hundreds of individual Escherichia coli cells over days, in a dynamically adjustable environment. This integrated platform broadly enables experiments that bridge individual and population behaviors. We demonstrate: (i) population structuring by independent closed-loop control of gene expression in many individual cells, (ii) cell–cell variation control during antibiotic perturbation, (iii) hybrid bio-digital circuits in single cells, and freely specifiable digital communication between individual bacteria. These examples showcase the potential for real-time integration of theoretical models with measurement and control of many individual cells to investigate and engineer microbial population behavior.

7.5. Balancing a genetic toggle switch by real-time feedback control and periodic forcing

Participants: Gregory Batt, Jean-Baptiste Lugagne, Melanie Kirch (former Member), Agnes Köhler (former Member), Sebastian Sosa Carrillo.

Cybergenetics is a novel field of research aiming at remotely pilot cellular processes in real-time with to leverage the biotechnological potential of synthetic biology. Yet, the control of only a small number of genetic circuits has been tested so far. Here we investigate the control of multistable gene regulatory networks, which are ubiquitously found in nature and play critical roles in cell differentiation and decision-making. Using an in silico feedback control loop, we demonstrate that a bistable genetic toggle switch can be dynamically maintained near its unstable equilibrium position for extended periods of time [2]. Importantly, we show that a direct method based on dual periodic forcing is sufficient to simultaneously maintain many cells in this undecided state. These findings pave the way for the control of more complex cell decision-making systems at both the single cell and the population levels, with vast fundamental and biotechnological applications.

7.6. Abstracting the dynamics of biological pathways using information theory: a case study of apoptosis pathway

Participants: Gregory Batt, François Bertaux (former Member), Sucheendra Palaniappan (former Member).

Quantitative models are increasingly used in systems biology. Usually, these quantitative models involve many molecular species and their associated reactions. When simulating a tissue with thousands of cells, using these large models becomes computationally and time limiting. In our paper, we propose to construct abstractions using information theory notions [3]. Entropy is used to discretize the state space and mutual information is used to select a subset of all original variables and their mutual dependencies. We apply our method to an hybrid model of TRAIL-induced apoptosis in HeLa cell. Our abstraction, represented as a Dynamic Bayesian Network (DBN), reduces the number of variables from 92 to 10, and accelerates numerical simulation by an order of magnitude, yet preserving essential features of cell death time distributions.

7.7. Long-term tracking of budding yeast cells in brightfield microscopy: CellStar and the Evaluation Platform

Participants: Gregory Batt, Artémis Llamosi (former Member).

With the continuous expansion of single cell biology, the observation of the behaviour of individual cells over extended durations and with high accuracy has become a problem of central importance. Surprisingly, even for yeast cells that have relatively regular shapes, no solution has been proposed that reaches the high quality required for long-term experiments for segmentation and tracking (S&T) based on brightfield images. In this contribution, we present CellStar, a tool chain designed to achieve good performance in long-term experiments [5]. The key features are the use of a new variant of parametrized active rays for segmentation, a neighbourhood-preserving criterion for tracking, and the use of an iterative approach that incrementally improves S&T quality. A graphical user interface enables manual corrections of S&T errors and their use for the automated correction of other, related errors and for parameter learning. We created a benchmark dataset with manually analysed images and compared CellStar with six other tools, showing its high performance, notably in long-term tracking. As a community effort, we set up a website, the Yeast Image Toolkit, with the benchmark and the Evaluation Platform to gather this and additional information provided by others.

7.8. Sensitivity estimation for stochastic models of biochemical reaction networks in the presence of extrinsic variability

Participant: Jakob Ruess.

Determining the sensitivity of certain system states or outputs to variations in parameters facilitates our understanding of the inner working of that system and is an essential design tool for the de novo construction of robust systems. In cell biology, the output of interest is often the response of a certain reaction network to some input (e.g., stressors or nutrients) and one aims to quantify the sensitivity of this response in the presence of parameter heterogeneity. We argue that for such applications, parametric sensitivities in their standard form do not paint a complete picture of a system's robustness since one assumes that all cells in the population have the same parameters and are perturbed in the same way. In the published contribution, we consider stochastic reaction networks in which the parameters are randomly distributed over the population and propose a new sensitivity index that captures the robustness of system outputs upon changes in the characteristics of the parameter distribution, rather than the parameters themselves [4]. Subsequently, we make use of Girsanov's likelihood ratio method to construct a Monte Carlo estimator of this sensitivity index. However, it turns out that this estimator has an exceedingly large variance. To overcome this problem, we propose a novel estimation algorithm that makes use of a marginalization of the path distribution of stochastic reaction networks and leads to Rao-Blackwellized estimators with reduced variance.

7.9. Recombinase-based genetic circuit optimization by integer linear programming

Participant: François Fages.

The rapid advancements of synthetic biology show promising potential in biomedical and other applications. Recently, recombinases were proposed as a tool to engineer genetic logic circuits with long-term memory in living and even mammalian cells. The technology is under active development, and the complexity of engineered genetic circuits grows continuously. However, how to minimize a genetic circuit composed of recombinase-based logic gates remain largely open. In [12], we formulate the problem as a cubic-time assignment problem and solved by a 0/1-ILP solver to minimize DNA sequence length of genetic circuits. Experimental results show effective reduction of our optimization method, which may be crucial to enable practical realization of complex genetic circuits.

7.10. Coupled models of the cell cycle and circadian clock

Participants: François Fages, Sylvain Soliman, Pauline Traynard (former Member).

Experimental observations have put in evidence autonomous self-sustained circadian oscillators in most mammalian cells, and proved the existence of molecular links between the circadian clock and the cell cycle. Some mathematical models have also been built to assess conditions of control of the cell cycle by the circadian clock, with applications to cancer chronotherapy optimization. However, recent studies in individual NIH3T3 fibroblasts have shown an unexpected acceleration of the circadian clock together with the cell cycle when the culture medium is enriched with growth factors, and the absence of such acceleration in confluent cells. In order to explain these observations , we have studied a possible entrainment of the circadian clock by the cell cycle through a regulation of clock genes around the mitosis phase. We developed a computational model in Biocham with a formal specification of the observed behavior in quantitative temporal logic to investigate the conditions of entrainment in period and phase. We showed that either the selective activation of RevErb-α or the selective inhibition of Bmal1 transcription during the mitosis phase, allowed us to fit the experimental data on both period and phase, while a uniform inhibition of transcription during mitosis seems incompatible with the phase data. In [6], we presented those results and some further predictions of the bidirectional model with a coupling in both directions.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR Projects

- ANR-MOST BIOPSY (2016-2020) on “Biochemical Programming System”, coordinated by F. Molina (CNRS, Sys2diag, Montpellier) and J.H. Jiang (National Taiwan University), with F. Fages.
- ANR MEMIP (2016-2020) on “Mixed-Effects Models of Intracellular Processes”, coordinated by G. Batt, with P. Hersen, (CNRS/Paris7), E. Cinquemani (Inria EPI IBIS) and M. Lavielle (Inria/CNRS/Polytechnique, EPI XPOP).
- ANR COGEX (2016-2019) on “Computer Aided Control of Gene Expression” coordinated by P. Hersen (MSC lab, CNRS/Paris7), with G. Batt and G. Truan (LISBP, CNRS/INSA).
- ANR Blanc HYCLOCK (2014-2018) on “Hybrid modeling of time for Circadian Clock Biology and Chronopharmacology”, coordinated by F. Delaunay (CNRS, Nice), with F. Lévi (INSERM Paris-Sud), G. Bernot (CNRS I3S, Nice), O. Roux (Ecole Centrale Nantes), F. Fages and S. Soliman.
- ANR Blanc STOCH-MC (2014-2018) on “Stochastic Models: Scalable Model Checking”, coordinated by Blaise Genest (Inria Rennes), with Grégory Batt, Wieslaw Zielonka (LIAFA), and Hugo Gimbert (LaBRI).

- ANR Investissement Avenir ICEBERG project (2011-2017) "From population models to model populations", coordinated by Grégory Batt, with Pascal Hersen (MSC lab, Paris Diderot Univ./CNRS), Reiner Veitia (Institut Jacques Monod, Paris Diderot Univ./CNRS), Olivier Gandrillon (BM2A lab, Lyon Univ./CNRS), Cédric Lhoussaine (LIFL/CNRS), and Jean Krivine (PPS lab, Paris Diderot Univ./CNRS).

8.1.2. Inria Project Lab

- IPL COSY (2017-2021) "real-time control of synthetic microbial communities", coordinated by Eugenio Cinquemani (Ibis, Inria), with Jean-Luc Gouzé (Biocore, Inria), Gregory Batt, Frédéric Bonnans (Commands, Inria), Efimov Denis (Non-A, Inria), and Hans Geiselmann (BIOP, Université Grenoble-Alpes), Beatrice Laroche (Maiage, Inra Jouy-en-Josas), and Hyun Youk (Youk lab, TU Delft).

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

- H2020 FET-OPEN COSY-BIO (2017-2020), "Control Engineering of Biological Systems for Reliable Synthetic Biology Applications", coordinated by Diego di Bernardo (Tigem), with Filippo Menolascina (Edinburgh U), Mario di Bernardo (Naples U), Pascal Hersen (Paris7 U), Mustafa Khammash (ETHZ), Gregory Batt, Guy-Bart Stan (Imperial College), and Lucia Marucci (Bristol U).

8.3. International Initiatives

8.3.1. Participation in International Programs

- French-German PROCOPE (2015-2017) grant on "Réduction de modèle et analyse de grands réseaux biochimiques par des méthodes stoechiométriques et tropicales", coord. Prof. Andreas Weber, University of Bonn, Germany, and Prof. Ovidiu Radulescu, Univ. Montpellier, France.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. Member of the Organizing Committees

François Fages is co-organizer of the CSBC Workshop on Computational Systems Biology for Cancer, Institut des Systèmes Complexes, Paris, France, 24-26 jan. 2018.

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

- François Fages was member of the PC of
  - CMSB'17 The 15th conference on Computational Methods for Systems Biology, September 27th to 29th 2017, Darmstadt, Germany.
  - FroCos'17 The 11th International Symposium on Frontiers of Combining Systems, 25-29 September 2017, Brasilia, Brazil.

9.1.2.2. Reviewer

- François Fages reviewed one article for the 15th Int. Computer Science Symposium in Russia CSR 2017.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

François Fages is member of

- the Editorial Board of the Computer Science area of the Royal Society Open Science journal since 2014,
- the Editorial Board of the journal RAIRO OR Operations Research since 2004.

9.1.3.2. Reviewer - Reviewing Activities

In addition to their Editorial Board and Program Committee duties,

- Grégory Batt reviewed one article for *Nature Scientific Reports*.
- François Fages reviewed articles for *Bioinformatics, ACM Transactions on Modeling and Computer Simulation, Fundamenta Informaticae, Computers and Industrial Engineering*
- Sylvain Soliman reviewed articles for *Briefings in Bioinformatics*, *PNAS*, *Transactions on Computational Biology and Bioinformatics*

9.1.4. Invited Talks

- Virgile Andréani gave a poster presentation at the Spring school on Computational Systems Biology (CompSysBio 2017), Aussois, March 19-25 2017
- Gregory Batt gave invited talks at
  - Computational biology seminar of Curie Institute, “Predicting long-term effects of repeated apoptosis-inducing drug treatments: coupling signal transduction pathways with stochastic protein turnover models”, Paris, Jan 19 2017.
  - 6th Annual Consortium Meeting of Revive LabEx, “Multi-scale modeling of TRAIL-induced apoptosis: from protein level fluctuations to drug resistance in multicellular spheroids”, Montvillargenne, Jan 24 2017.
  - ENS/ESPCI joint Biophysics seminar, “Balancing a genetic toggle switch by real-time control or periodic stimulations”, Paris, Feb 10 2017.
  - BioNetVisA 2017 workshop “From biological network reconstruction to data visualization and analysis in molecular biology and medicine”, “A multi-scale model for investigating TRAIL resistance in multi-cellular tumor spheroids”, Basel, Sept 12 2017.
  - Workshop on Control of Cellular and Molecular Systems, “Balancing a genetic toggle switch by real-time control or periodic stimulations”, Mathematical Biosciences Institute, Ohio State University.
  - Second Symposium du Réseau de Biologie des Systèmes de Sorbonne Université, “Balancing a genetic toggle switch by real-time control or periodic stimulations”, Paris, Dec 1 2017.
- François Fages gave invited talks at
  - National Taiwan University, Dept. Electrical Engineering, “Strong Turing Completeness of Continuous Chemical Reaction Networks and Compilation of Mixed Analog-Digital Programs”, Tapei, Taiwan, 9 Nov 2017.
  - Centre de Recherche Interdisciplinaire, CRI Paris, “Living with Algorithms and Algorithms of the Living”, Paris, 29 June 2017.
  - GT BIOSS-preGDR IA, keynote talk “In Quest of the Software of the Living: Successes and Difficulties of the Program Verification Paradigm in Cell Biology”, Gif sur Yvette, 22 June 2017.
  - Dynamic Days Europe, “Quantitative Modelling of the Circadian Clock and Cell Cycle Coupling: Prediction of RevErb- alpha Up-Regulation during Mitosis in Mouse Embryonic Fibroblasts – Temporal Logic Constraints and Parameter Search in BIOCHAM”, Szeged, Hungary, 7 June 2017.
  - “Turing Completeness of Biochemical Reactions over a finite set of molecules under the Differential Semantics and Compilation of Mixed Analog-Digital Programs”, French Ambassy in Berlin, Germnany, 1 June 2017.
  - GT BIOSS “Turing Completeness of Biochemical Reactions under the Continuous Semantics and Compilation of Mixed Digital-Analog Programs”, Montpellier, 13 March 2017.

  - Lab Sys2diag “Biochemical Compiler and Mixed Analog-Digital Algorithms in the Cell”, Montpellier, 9 March 2017.
- Jean-Baptiste Lugagne gave invited talks at
  - Khammash lab seminar, “Balancing a genetic toggle switch by real-time feedback control and periodic forcing”, ETHZ, Zurich, March 2017
  - Seminar of SynthSys Centre for Synthetic and Systems Biology, “Balancing a genetic toggle switch by real-time feedback control and periodic forcing”, University of Edinburgh, June 08 2017
  - International Workshop on Control Engineering and Synthetic Biology, ‘Balancing a genetic toggle switch by real-time feedback control and periodic forcing” (poster presentation), Royal Academy of Engineering, London, July 17 2017
  - Seminar of Theory of Living Matter Group, “Balancing a genetic toggle switch by real-time feedback control and periodic forcing”, University of Cambridge, July 19 2017
- Jakob Ruess gave invited talks at
  - Working days of GT BIOSS “Control of bio-digital systems in single cells”, March 13 2017
  - Conference of GdR Biologie Synthétique et Systémique (BioSynSys) “Control of bio-digital systems in single cells”, La Grande Motte, Oct 10 2017

9.1.5. Leadership within the Scientific Community

- Grégory Batt is a member of
  - the IEEE/CSS Technical Committee on Systems Biology,
  - the scientific board of the GDR de Biologie de Synthèse et des Systèmes
  - the GDR de Bioinformatique Moléculaire, in charge of the axis on Biological network modelling, systems biology and synthetic biology
  - co-animator of the French working group on Symbolic Systems Biology GT BIOSS
  - the scientific committee of the Spring school on Computational Systems Biology (CompSysBio 2017)
- François Fages is a member of
  - the Steering Committee of the International Conference on Computational Methods for Systems Biology since 2008,
  - the Scientific Council of the *Doctorate School “Frontières Du Vivant”* at *Center for Research and Interdisciplanirity*, Universities Paris Descartes and Paris Diderot, since 2010,
  - The Scientific Commmittee of the Summer School Ecole Thématique Modélisation Formelle des Réseaux de Régulation Biologique since 2010.

9.1.6. Scientific Expertise

- Gregory Batt
  - has been in charge of the definition of the 16th challenge of the 2018-2022 Inria strategic plan on predictive systems biology
  - has been a member of the INRA selection committee for hiring permanent junior researchers (CR2)
  - has evaluated the PhD thesis of Lorena Postiglioni and Giansimone Perrino (Diego di Bernardo, Tigem, Italy; Thesis evaluation report)

  - has been a member of thesis advisory committee meetings for Arnaud Bonnaffoux (Olivier Gandrillon ENS Lyon/Cosmo Company), Raphael Goujet (Ariel Lindner, CRI, Paris), Adele Kerjouan (Olivier Destaing, IAB, Grenoble), and Mathilde Koch (Jean-Loup Faulon, INRA, Jouy en Josas)
- François Fages
  - is a member of the jury for the *Prix de thèse Gilles Kahn* of the *Société Informatique de France*, since 2015,
  - reviewed one Research Fellowship application for the Royal Commission for the Exhibition of 1851

9.1.7. Research Administration

- François Fages is member of the "Bureau du Comité des Projets" of Inria Saclay-IdF.
- Sylvain Soliman is member of the "Commission Scientifique" of Inria Saclay-IdF

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master: Grégory Batt (coordinator and teacher: 48h) and Jakob Ruess (24h), *Computational Biology*, M1, Master Approches Interdisciplinaires du Vivant (AIV).

Master/PhD: Grégory Batt (co-coordinator 80h, teacher 8h) and Jakob Ruess (8h), *Modeling and engineering of biological systems*, M2/PhD, Institut de Technologie et d'Innovation of Paris Sciences et Lettres (PSL-ITI), Paris.

Master/PhD: Gregory Batt (Scientific committee and teacher 6h), Spring school on Computational Systems Biology (CompSysBio 2017), Aussois, March 19-25 2017

Master: François Fages (coordinator module 48h, teaching 12h), C2-19 *Computational Methods for Systemic and Synthetic Biology*, Master Parisien de Recherche en Informatique (MPRI), Paris.

Master: Chiara Fracassi (20h), *Experimental Methods in Biophysics*, M1, Master Approches Interdisciplinaires du Vivant (AIV).

Master: Sylvain Soliman, C2-35-1 *Constraint Programming*, coordinator and teaching 24h, M2, Master Parisien de Recherche en Informatique (MPRI), Paris.

Master: Denis Thieffry (coordinator, 6h) and Gregory Batt (6h), *Dynamical Modelling of Cellular Regulatory Networks*, M2, Interdisciplinary Master in Life Science at the Ecole Normale Supérieure, Paris.

E-learning

Online BIOCHAM tutorial notebook by F. Fages and S. Soliman, presented at CMSB 2017, Darmstadt, Germany.

9.2.2. Supervision

Grégory Batt is currently supervising the Ph.D. theses of

- Virgile Andreani (ED STIC, ENS)
- Sebastian Sosa Carrillo (ED FdV, Inria)
- Chetan Aditya (ED FdV, Inria)

9.2.3. Juries

- Ph.D.: Stefano Casagranda (Inria Sophia Antipolis), *G. Batt*, June 30 2017.
- HDR: Russ Harmer, Ecole Normale Supérieure de Lyon, *F. Fages, reviewer*, May 22 2017.
- Ph.D.: Guillaume Madelaine, University of Lille, *F. Fages, reviewer*, Feb. 28 2017.
- Ph.D.: Danhua Peng, University of Rostock, Germany, *F. Fages, reviewer*, Feb. 3 2017.
- Ph.D.: Ferdinanda Camporesi, University PSL, ENS Paris, *F. Fages*, Jan 23 2017.

9.3. Popularization

We participated in seminars for general audiences.

- G. Batt, *"La cybergénétique: piloter des cellules par ordinateur"*, 50 ans d'Inria au Siège, Rocquencourt, Nov 16 2017.
- F. Fages, *"Living with algorithms and algorithms of the living"*, Workshop Shaping Major Transitions in the Digital World. Interdisciplinary Research Center, CRI, Paris, 28-30 June 2017.
- J. Ruess, *"Automating scientific discovery in single cells"*, Unithé ou café, Inria Saclay - Ile de France, Apr 21 2017

Our publications also raised the attention of specialized and non-specialized media. One can mention

- "Piloter le comportement de cellules biologiques par ordinateur, c'est possible" in Science et Avenir
- "Forscher verbinden und kontrollieren Bakterien mit einem Computer", in Wiener Zeitung
- "Forscher steuern Bakterien durch Verbindung mit Computer" in der Standard
- "Forscher verbinden und kontrollieren Bakterien mit einem Computer " and "Virtuelle Realität für Bakterien" in Austria Presse Agentur (APA)
- "Ils parviennent à contrôler des cellules biologiques par ordinateur" in SciencePost
- "Computerized biology, or how to control a population of cells with a computer" and "Virtual reality for bacteria" in Phys.org
- "Computerized biology, or how to control a population of cells with a computer" in EurekAlert.org
- "Scientists Learn to Control Bacteria by Connecting Them to Computers" in Scicasts.com
- "Virtuelle Realität für Bakterien" in Innovations-report.de
- "Virtuelle Realität für Bakterien" in Informationsdienst Wissenschaft

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] R. Chait, J. Ruess, T. Bergmiller, G. Tkačik, C. C. Guet.*Shaping bacterial population behavior through computer-interfaced control of individual cells*, in "Nature Communications", November 2017, vol. 8, n^o^ 1, 1535 [*DOI :* 10.1038/s41467-017-01683-1], https://hal.inria.fr/hal-01669883.

[2] J.-B. Lugagne, S. Sosa Carrillo, M. Kirch, A. Köhler, G. Batt, P. Hersen.*Balancing a genetic toggle switch by real-time feedback control and periodic forcing*, in "Nature Communications", December 2017, vol. 8, n^o^ 1, p. 1-8 [*DOI :* 10.1038/s41467-017-01498-0], https://hal.inria.fr/hal-01666980.

[3] S. K. PALANIAPPAN, F. BERTAUX, M. PICHENÉ, E. FABRE, G. BATT, B. GENEST.*Abstracting the dynamics of biological pathways using information theory: a case study of apoptosis pathway*, in "Bioinformatics", February 2017, vol. 33, n^{o} 13, p. 1980 - 1986 [*DOI :* 10.1093/BIOINFORMATICS/BTX095], https://hal.inria.fr/hal-01547618.

[4] J. RUESS, H. KOEPPL, C. ZECHNER.*Sensitivity estimation for stochastic models of biochemical reaction networks in the presence of extrinsic variability*, in "Journal of Chemical Physics", 2017, vol. 146, n^{o} 12, 124122 [*DOI :* 10.1063/1.4978940], https://hal.archives-ouvertes.fr/hal-01547484.

[5] C. VERSARI, S. STOMA, K. BATMANOV, A. LLAMOSI, F. MROZ, A. KACZMAREK, M. DEYELL, C. LHOUSSAINE, P. HERSEN, G. BATT.*Long-term tracking of budding yeast cells in brightfield microscopy: CellStar and the Evaluation Platform*, in "Interface", 2017, vol. 14, n^{o} 127, 32 [*DOI :* 10.1098/RSIF.2016.0705], https://hal.inria.fr/hal-01547623.

Invited Conferences

[6] F. FAGES.*Model-based Investigation of the Circadian Clock and Cell Cycle Coupling in Mouse Embryonic Fibroblasts: Prediction of RevErb-α Up-regulation During Mitosis*, in "Dynamics Days 2017 - XXXVII Dynamics Days Europe International Conference", Szeged, Hungary, June 2017, https://hal.inria.fr/hal-01651833.

International Conferences with Proceedings

[7] A. CARCANO, F. FAGES, S. SOLIMAN.*Probably Approximately Correct Learning of Regulatory Networks from Time-Series Data*, in "CMSB'17 - 15th International Conference on Computational Methods for Systems Biology", Darmstadt, Germany, J. FERET, H. KOEPPL (editors), September 2017, vol. Lecture Notes in Computer Science, n^{o} 10545, p. 74-90, https://hal.inria.fr/hal-01519826.

[8] *Best Paper*
F. FAGES, G. LE GULUDEC, O. BOURNEZ, A. POULY.*Strong Turing Completeness of Continuous Chemical Reaction Networks and Compilation of Mixed Analog-Digital Programs*, in "CMSB 2017 - 15th International Conference on Computational Methods in Systems Biology", Darmstadt, Germany, J. FERET, H. KOEPPL (editors), Proceedings of the fiveteen international conference on Computational Methods in Systems Biology, CMSB 2017, September 2017, vol. Lecture Notes in Computer Science, n^{o} 10545, p. 108-127, https://hal.inria.fr/hal-01519828.

Scientific Books (or Scientific Book chapters)

[9] F. BERTAUX, D. DRASDO, G. BATT.*System modeling of receptor-induced apoptosis*, in "TRAIL, Fas Ligand, TNF and TLR3 in Cancer", O. MICHEAU (editor), Resistance to Targeted Anti-Cancer Therapeutics, 2017, n^{o} 12, https://arxiv.org/abs/1712.06822 , https://hal.inria.fr/hal-01667015.

Scientific Popularization

[10] F. FAGES, G. LE GULUDEC.*Programmes biochimiques et algorithmes mixtes analogiques-digitaux dans la cellule*, in "Sciences de la vie, sciences de l'information", T. GAUDIN, D. LACROIX, M.-C. MAUREL, J.-C. POMEROL (editors), Colloque de Cérisy, ISTE editions, September 2017, https://hal.inria.fr/hal-01409743.

Other Publications

[11] F. FAGES, T. MARTINEZ, D. A. ROSENBLUETH, S. SOLIMAN.*Influence Networks compared with Reaction Networks: Semantics, Expressivity and Attractors*, October 2017, Submitted to IEEE/ACM TRANSACTIONS ON COMPUTATIONAL BIOLOGY AND BIOINFORMATICS, https://hal.inria.fr/hal-01510216.

[12] C.-N. LAI, J.-H. JIANG, F. FAGES.*Recombinase-Based Genetic Circuit Optimization*, October 2017, p. 1-4, BioCAS 2017 - 13th IEEE Biomedical Circuits and Systems Conference, Poster, https://hal.inria.fr/hal-01659183.

Project-Team M3DISIM

Mathematical and Mechanical Modeling with Data Interaction in Simulations for Medicine

IN COLLABORATION WITH: Laboratoire de Mécanique des Solides

IN PARTNERSHIP WITH:
Ecole Polytechnique

RESEARCH CENTER
Saclay - Île-de-France

THEME
Modeling and Control for Life Sciences

Table of contents

Project-Team M3DISIM

Creation of the Team: 2013 January 01, updated into Project-Team: 2016 June 01

Keywords:

Computer Science and Digital Science:

A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.2. - Stochastic Modeling (SPDE, SDE)
A6.1.4. - Multiscale modeling
A6.1.5. - Multiphysics modeling
A6.2.1. - Numerical analysis of PDE and ODE
A6.3.1. - Inverse problems
A6.3.2. - Data assimilation
A6.3.4. - Model reduction
A6.4.1. - Deterministic control
A6.4.3. - Observability and Controlability
A6.4.4. - Stability and Stabilization

Other Research Topics and Application Domains:

B1.1.10. - Mathematical biology
B2.2.1. - Cardiovascular and respiratory diseases
B2.6.2. - Cardiac imaging
B2.6.3. - Biological Imaging

1. Personnel

Research Scientists

Philippe Moireau [Team leader, Inria, Senior Researcher, HDR]
Radomir Chabiniok [Inria, Starting Research Position]
Dominique Chapelle [Inria, Senior Researcher, HDR]
Sébastien Imperiale [Inria, Researcher]

Faculty Members

Jean-Marc Allain [Ecole polytechnique, Associate Professor, HDR]
Martin Genet [Ecole polytechnique, Associate Professor]
Patrick Le Tallec [Ecole polytechnique, Professor, HDR]

External Collaborators

Fabrice Vallee [Assistance publique/Hôpitaux de Paris, Researcher]
Matthieu Caruel [Univ Paris-Val de Marne]
Alexandre Laurin [Independent, from Oct 2017]
Hajer Methenni [CEA, from Oct 2017]

Technical Staff

Alexandre Laurin [Inria, until Jul 2017]

PhD Students

Ezgi Berberoglu [ETH Zurich, PhD Student, from May 2017]
Federica Caforio [Inria]
Marija Gusseva [Inria, from Oct 2017]

Ustim Khristenko [Ecole polytechnique]
François Kimmig [Ecole polytechnique]
Arthur Le Gall [Assistance publique/Hôpitaux de Paris]
Cecile Patte [Inria, from Nov 2017]
Florent Wijanto [Ecole polytechnique]

Post-Doctoral Fellow
Jean-Sébastien Affagard [Ecole polytechnique]

Visiting Scientists
Jorge Albella Martinez [USC Santiago de Compostela, from Sep 2017]
Peter Baumgartner [Ecole polytechnique, from Sep 2017]
Nicole Tueni [Ecole polytechnique, from Sep 2017]

Administrative Assistant
Hélèna Kutniak [Inria]

2. Overall Objectives

2.1. Overall Objectives

The research carried out in the M3DISIM team has a rather global methodological perspective oriented towards biomechanics, encompassing mathematical modeling and analysis, inverse problems arising from model-data coupling, and the formulation and analysis of effective and reliable numerical procedures adapted to this overall program. We are also very keen on demonstrating the effectiveness and relevance of these methods in actual applications, usually by proof-of-concept studies carried out within various collaborations.

3. Research Program

3.1. Multi-scale modeling and coupling mechanisms for biomechanical systems, with mathematical and numerical analysis

Over the past decade, we have laid out the foundations of a multi-scale 3D model of the cardiac mechanical contraction responding to electrical activation. Several collaborations have been crucial in this enterprise, see below references. By integrating this formulation with adapted numerical methods, we are now able to represent the whole organ behavior in interaction with the blood during complete heart beats. This subject was our first achievement to combine a deep understanding of the underlying physics and physiology and our constant concern of proposing well-posed mathematical formulations and adequate numerical discretizations. In fact, we have shown that our model satisfies the essential thermo-mechanical laws, and in particular the energy balance, and proposed compatible numerical schemes that – in consequence – can be rigorously analyzed, see [4]. In the same spirit, we have recently formulated a poromechanical model adapted to the blood perfusion in the heart, hence precisely taking into account the large deformation of the mechanical medium, the fluid inertia and moving domain, and so that the energy balance between fluid and solid is fulfilled from the model construction to its discretization, see [5].

3.2. Inverse problems with actual data – Fundamental formulation, mathematical analysis and applications

A major challenge in the context of biomechanical modeling – and more generally in modeling for life sciences – lies in using the large amount of data available on the system to circumvent the lack of absolute modeling ground truth, since every system considered is in fact patient-specific, with possibly non-standard conditions associated with a disease. We have already developed original strategies for solving this particular type of inverse problems by adopting the observer stand-point. The idea we proposed consists in incorporating to the classical discretization of the mechanical system an estimator filter that can use the data to improve the quality of the global approximation, and concurrently identify some uncertain parameters possibly related to a diseased state of the patient, see [7], [8], [9]. Therefore, our strategy leads to a coupled model-data system solved similarly to a usual PDE-based model, with a computational cost directly comparable to classical Galerkin approximations. We have already worked on the formulation, the mathematical and numerical analysis of the resulting system – see [3] – and the demonstration of the capabilities of this approach in the context of identification of constitutive parameters for a heart model with real data, including medical imaging, see [1].

4. Application Domains

4.1. Clinical applications

After several validation steps – based on clinical and experimental data – we have reached the point of having validated the heart model in a pre-clinical context where we have combined direct and inverse modeling in order to bring predictive answers on specific patient states. For example, we have demonstrated the predictive ability of our model to set up pacemaker devices for a specific patient in cardiac resynchronization therapies, see [11]. We have also used our parametric estimation procedure to provide a quantitative characterization of an infarct in a clinical experiment performed with pigs, see [1].

5. Highlights of the Year

5.1. Highlights of the Year

- Promotion of Jean-Marc Allain as a professor at Polytechnique.
- Patent submitted and accepted on heart and vessels modelling with data interaction ([40]).
- Submission of a IHU proposal, of 3 ERC proposals, 1 associated team proposal with UT Southwestern Medical Center Dallas
- Contract of collaboration with UT Southwestern Medical Center Dallas (Profs. G. Greil and T. Hussain)

6. New Software and Platforms

6.1. FELiScE

Finite Elements for Life SCiences and Engineering problems
KEYWORDS: Finite element modelling - Cardiac Electrophysiology - Cardiovascular and respiratory systems

FUNCTIONAL DESCRIPTION: FELiScE is a finite element code which the M3DISIM and REO project-teams have decided to jointly develop in order to build up on their respective experiences concerning finite element simulations. One specific objective of this code is to provide in a unified software environment all the state-of-the-art tools needed to perform simulations of the complex respiratory and cardiovascular models considered in the two teams – namely involving fluid and solid mechanics, electrophysiology, and the various associated coupling phenomena. FELISCE is written in C++, and may be later released as an opensource library. FELiScE was registered in July 2014 at the Agence pour la Protection des Programmes under the Inter Deposit Digital Number IDDN.FR.001.350015.000.S.P.2014.000.10000.

- Participants: Axel Fourmont, Benoit Fabreges, Damiano Lombardi, Dominique Chapelle, Faisal Amlani, Irène Vignon-Clementel, Jean-Frédéric Gerbeau, Marina Vidrascu, Matteo Aletti, Miguel Angel Fernandez Varela, Mikel Landajuela Larma, Philippe Moireau and Sébastien Gilles
- Contact: Jean-Frédéric Gerbeau
- URL: http://felisce.gforge.inria.fr

6.2. HeartLab

KEYWORDS: Computational geometry - Image analysis - Cardiac - Health - Simulation
SCIENTIFIC DESCRIPTION: The heartLab software is a library designed to perform both simulation and estimation of the heart mechanical behavior (based on various types of measurements, e.g. images).

Also included are geometric data and tools in the code to define cardiac anatomical models compatible with the simulation requirements in terms of mesh quality, fiber direction data defined within each element, and the referencing necessary for handling boundary conditions and estimation, in particular. These geometries are analytical or come from computerized tomography (CT) or magnetic resonance (MR) image data of humans or animals.
FUNCTIONAL DESCRIPTION: The heartLab software is a library designed to perform both simulation and estimation of the heart mechanical behavior (based on various types of measurements, e.g. images).

Also included are geometric data and tools in the code to define cardiac anatomical models compatible with the simulation requirements in terms of mesh quality, fiber direction data defined within each element, and the referencing necessary for handling boundary conditions and estimation, in particular. These geometries are analytical or come from computerized tomography (CT) or magnetic resonance (MR) image data of humans or animals.

- Participants: Radomir Chabiniok, Gautier Bureau, Martin Genet, Federica Caforio, Ustim Khristenko, Dominique Chapelle and Philippe Moireau
- Contact: Philippe Moireau
- URL: https://raweb.inria.fr/rapportsactivite/RA2013/m3disim/uid14.html

6.3. Verdandi

KEYWORDS: HPC - Model - Software Components - Partial differential equation
FUNCTIONAL DESCRIPTION: Verdandi is a free and open-source (LGPL) library for data assimilation. It includes various such methods for coupling one or several numerical models and observational data. Mainly targeted at large systems arising from the discretization of partial differential equations, the library is devised as generic, which allows for applications in a wide range of problems (biology and medicine, environment, image processing, etc.). Verdandi also includes tools to ease the application of data assimilation, in particular in the management of observations or for a priori uncertainty quantification. Implemented in C++, the library may be used with models implemented in Fortran, C, C++ or Python.

- Participants: Dominique Chapelle, Gautier Bureau, Nicolas Claude, Philippe Moireau and Vivien Mallet
- Contact: Vivien Mallet
- URL: http://verdandi.gforge.inria.fr/

6.4. CardiacLab

KEYWORDS: Cardiovascular and respiratory systems - Matlab - Real time

FUNCTIONAL DESCRIPTION: CardiacLab is a MATLAB toolbox allowing to perform "real-time" cardiac simulations using 0D models of the cardiovascular systems. Its modular development includes (1) a module integrating the mechanical dynamics of the cavity taking into account its particular geometry, (2) a module allowing to choose a micro-model of the cardiac contraction, (3) a module of phase management, (4) a circulation module based on Windkessel models or more advanced 1D flows models, and (5) a perfusion module. The objective of this code is threefold: (1) demonstrate to students, engineers, medical doctors, the interest of modeling in cardiac applications, (2) unify our original modeling developments with the possibility to evaluate them with previous team developments before integrating them into 3D complex formulations, and (3) explore some avenues pertaining to real-time simulat

- Participants: Sebastien Impériale, Martin Genet, Federica Caforio, Ustim Khristenko, Peter Baumgartner, Radomir Chabiniok, François Kimmig and Arthur Le Gall
- Contact: Philippe Moireau
- URL: https://gitlab.inria.fr/M3DISIM/CardiacLab

7. New Results

7.1. Mathematical and Mechanical Modeling

7.1.1. Modelling of collagen fibers elastic properties

Participants: Peter Baumgartner, Florent Wijanto, Jean-Marc Allain [correspondant], Matthieu Caruel [Univ. Paris-Est].

Our studies on collagen tissues have shown that the collagen fibers are able to elongate inelastically under stretch. In tendons, this effect has been attributed to the non-permanent cross-bridges that connect the different collagen fibrils (to assemble a fiber). This sliding effect appears experimentally to be reversible (at least partially) if the tissue is left long enough at its initial resting length. However, this sliding is classically included as an irreversible plastic response, or as a damage of the tissue. We are building a model based on a stochastic description of the binding and unbinding of the cross-bridges. This approach will enable us to have a microscopically based picture of the sliding, which will be able to explain some alterations in case of aging or pathologies of the tissue. At the moment, we have shown the importance of the density of cross-bridges in the cooperative response of the system. A publication is in preparation on the topic.

7.1.2. Multi-scale modeling of cardiac contraction

Participants: François Kimmig, Matthieu Caruel [Univ. Paris-Est], Dominique Chapelle [correspondant], Philippe Moireau.

This work aims at proposing a set of models of the muscular contraction targeting different scales in time and space and that can be used in the context of heart simulation. To this end, we developed so far two models using different approaches for the modeling of the force generating process at the molecular level called power stroke. First, we revised the standard chemo-mechanical models, which see the power stroke as a series of chemical states. Following the idea introduced by Truskinovsky and collaborators describing the power stroke as a continuum of mechanical states with the dynamics of the myosin head in the prescribed energy landscape governed by Langevin equations, we incorporated the attachment and detachment dynamics in the form of jump processes. In a second step, noting that the power stroke time scale is much shorter than that of heart contraction, we eliminated the power strokes dynamics and derived a two state – attached and detached – simplified model, each state being in fact associated with a statistical distribution of myosin head configurations. Both models have been integrated into our simulation framework CardiacLab, in order to benefit from the other modeling compartments available in the code, such as the geometrically reduced model of the heart left ventricle also developed in the team. These modeling elements will be confronted with experiments performed on cardiac muscle cells by collaborators in the team of Professor Lombardi at the University of Florence.

7.1.3. Mathematical and numerical modeling of shear waves propagation in the heart

Participants: Federica Caforio, Sébastien Imperiale [correspondant], Dominique Chapelle.

Shear acoustic waves remotely induced by the acoustic radiation force (ARF) of a focused ultrasound beam generated by piezoelectric sensors have been recently used in biomedical applications, e.g. in transient elastography techniques. By measuring the propagation velocity of generated shear waves in biological tissues, it is possible to locally assess biomechanical properties highly sensitive to structural changes corresponding to physiological and pathological processes. Recent experimental studies show the feasibility of applying transient elastography to the cardiac setting. In this context, the wave propagation induced by the ARF is superposed with the nonlinear mechanics associated with the heart deformation during the cardiac cycle. The aim of this work is to mathematically justify an original expression of the excitation induced by the ARF in nonlinear solids, based on energy considerations and asymptotic analysis. In soft media the propagation velocity of shear waves ($1 - -10\mathrm{m.s}^{-1}$) is much smaller than the velocity of pressure waves ($1500\mathrm{m.s}^{-1}$). The approach we propose consists in considering a family of problems parametrised by a small parameter ε related to the velocity ratio between the two wave propagation phenomena, the high frequency of the piezoelectric source term and the viscosity. In order to derive a simplified model for the expression of ARF, we investigate the limit behavior of the solution for $\varepsilon \to 0$. We show that the leading term of the expansion is related to the underlying nonlinear mechanics of the heart deformation, and the first two correction terms correspond to a fast-oscillating pressure wave excited by the probe, and an elastic field having as source term a nonlinear function of the first corrector. This field corresponds to the shear acoustic wave induced by the ARF.

7.1.4. Analysis and 2-scale convergence of a heterogeneous microscopic bidomain model

Participants: Sébastien Imperiale [correspondant], Annabelle Collin [Monc].

The aim of this work is to provide a complete mathematical analysis of the periodic homogenization procedure that leads to the macroscopic bidomain model in cardiac electrophysiology. We consider space-dependent and tensorial electric conductivities as well as space-dependent physiological and phenomenological non-linear ionic models. We provide the nondimensionalization of the bidomain equations and derive uniform estimates of the solutions. The homogenization procedure is done using 2-scale convergence theory which enables us to study the behavior of the non-linear ionic models in the homogenization process.

7.1.5. A reduced thoracic model for inverse problem solving in seismocardiography

Participants: Alexandre Laurin, Sébastien Imperiale [correspondant], Dominique Chapelle, Philippe Moireau.

Seismocardiography (SCG) is the study of low-frequency (< 60 Hz) vibrations of the thorax caused by the beating heart. Although it is assumed that SCG signals are caused by forces applied on the interior of the thorax by the heart, no comprehensive model exists to describe the parameters and relationships that govern the system. The main goal of this study is to describe in some detail the filter applied by the thorax on cardiac forces, taking into account its zone of contact with the heart as well as the zone of measurement, i.e. the location of the accelerometer on a participant's chest. A secondary goal is to identify the smallest set of parameters capable of reproducing the filter, reducing the model while retaining its capacity to lend itself to physiological interpretation. Finally, we described a method to use the reduced model to estimate cardiac forces from measured thoracic accelerations. The overall aim of the study is to develop numerical methods that can augment the existing SCG interpretations to include mechanical indices of the heart muscle, and do so in real time.

7.2. Numerical Methods

7.2.1. Numerical methods for computing cyclic steady states

Participants: Ustim Khristenko, Patrick Le Tallec [correspondant].

This work is focused on two techniques for fast computing of the steady cyclic states of evolution problems in non-linear mechanics with space-time periodicity conditions. This kind of problems can be faced in various applications, for instance in the rolling of a tyre with periodic sculptures as well as in a beating heart. Direct solvers for such problems are not very convenient, since they require the inversion of very large matrices. In order to avoid this, a cyclic solution is usually computed as an asymptotic limit of the associated initial value problem with prescribed initial data. However, when the relaxation time is high, convergence to the limit cycle can be very slow. The first technique considered is the Newton-Krylov method, looking for the unknown initial state that provides the space-time periodic solution. This initial state is defined by the space-time periodicity condition, solved with the Newton-Raphson technique. Since the associated Jacobian cannot be expressed explicitly, the method uses one of the matrix-free Krylov iterative solvers. Using information stored while computing the residual to solve the linear system makes its calculation time negligible with respect to the residual calculation time. The second method is the delayed feedback control: an observer-controller type modification of the standard evolution to the limit cycle by introducing a feedback control term, based on the periodicity error. The main result is the optimal form of the control term for a very general class of linear evolution problems, providing the fastest convergence to the cyclic solution. This control has also been adapted and tested for nonlinear problems. The methods discussed have been assessed using academic applications and they have also been implemented into the Michelin industrial code – applied to the rolling tyre model – as well as into the M3DISIM code for the cardiac contraction problem.

7.2.2. *Solving isotropic elastodynamics using potentials*

Participants: Sébastien Imperiale [correspondant], Jorge Albella.

This work has the potential to provide an original efficient method for the computations of elastic waves propagation in soft media (such as biological tissues), based on the property that pressure and shear waves decouple in isotropic media. Towards this direction, we considered the numerical solution of 2D elastodynamics isotropic equations using the decomposition of the displacement fields into potentials. This appears as a challenge for finite element methods, and we have addressed here the particular question of free boundary conditions. A stable (mixed) variational formulation of the evolution problem is proposed.

7.2.3. *The Arlequin method for transient wave scattering by small obstacles*

Participants: Sébastien Imperiale [correspondant], Jorge Albella.

In this work we extend the Arlequin method to overlapping domain decomposition technique for transient wave equation scattering by small obstacles. The main contribution of this work is to construct and analyze some variants of the Arlequin method from the continuous level to the fully discrete level. The constructed discretizations allow to solve wave propagation problems while using non-conforming and overlapping meshes for the background propagating medium and the surrounding of the obstacle, respectively. Hence we obtain a flexible and stable method in terms of the space discretization – an inf-sup condition is proven – while the stability of the time discretization is ensured by energy identities.

7.2.4. *Construction of a fourth-order time scheme for dissipative wave equations*

Participants: Sébastien Imperiale [correspondant], Juliette Chabassier [Magique-3d], Julien Diaz [Magique-3d].

This works deals with the construction of a fourth-order, energy preserving, explicit time discretization for dissipative linear wave equations. This discretization is obtained by replacing the inversion of a matrix – that comes naturally after using the technique of the Modified Equation on the second order Leap Frog scheme applied to dissipative linear wave equations – by an explicit approximation of its inverse. The stability of the scheme is studied first using an energy analysis, then an eigenvalue analysis. Numerical results in 1D illustrate the good behavior regarding space/time convergence and the efficiency of the newly derived scheme compared to more classical time discretizations. A loss of accuracy is observed for non-smooth profiles of dissipation, and we propose an extension of the method that fixes this issue. Finally, we assess the good performance of the scheme for a realistic dissipation phenomenon in Lorentz materials.

7.2.5. Coupled variational formulations of linear elasticity and the DPG methodology

Participant: Patrick Le Tallec [correspondant].

In this work, we develop a general approach akin to domain-decomposition methods to solve a single linear PDE, but where each subdomain of a partitioned domain is associated with a distinct variational formulation coming from a mutually well-posed family of so-called broken variational formulations of the original PDE. It can be exploited to solve challenging problems in a variety of physical scenarios where stability or a particular mode of convergence is desired in some part of the domain. The linear elasticity equations are solved in this work, but the approach can be applied to other equations as well. The broken variational formulations, which are essentially extensions of more standard formulations, are characterized by the presence of mesh-dependent broken test spaces and interface trial variables at the boundaries of the elements of the mesh. This allows necessary information to be naturally transmitted between adjacent subdomains, resulting in coupled variational formulations which are then proved to be globally well-posed. They are solved numerically using the DPG methodology, which is especially crafted to produce stable discretizations of broken formulations. Finally, expected convergence rates are verified in two different illustrative examples. This work has resulted in the publication [19].

7.2.6. A discontinuous Galerkin approach for cardiac electrophysiology

Participant: Radomir Chabiniok [correspondant].

Cardiac electrophysiology simulations are numerically challenging due to the propagation of a steep electrochemical wave front, and thus require discretizations with small mesh sizes to obtain accurate results. In this work – in collaboration with the Institute for Computational Mechanics, Technical University Munich and published in [21] – we present an approach based on the Hybridizable Discontinuous Galerkin method (HDG), which allows an efficient implementation of high-order discretizations into a computational framework. In particular using the advantage of the discontinuous function space, we present an efficient p-adaptive strategy for accurately tracking the wave front. HDG allows to reduce the overall degrees of freedom in the final linear system to those only on the element interfaces. Additionally, we propose a rule for a suitable integration accuracy for the ionic current term depending on the polynomial order and the cell model to handle high-order polynomials. Our results show that for the same number of degrees of freedom coarse high-order elements provide more accurate results than fine low-order elements. Introducing p-adaptivity further reduces computational costs while maintaining accuracy by restricting the use of high-order elements to resolve the wave front. For a patient-specific simulation of a cardiac cycle, p-adaptivity reduces the average number of degrees of freedom by 95% compared to the non-adaptive model. In addition to reducing computational costs, using coarse meshes with our p-adaptive high-order HDG method also simplifies practical aspects of mesh generation and postprocessing.

7.3. Inverse Problems

7.3.1. Discrete-time optimal filtering or Mortensen observer discretization

Participant: Philippe Moireau [correspondant].

In this work, we seek exact formulations of the optimal estimator and filter for a non-linear framework, as the Kalman filter is for a linear framework. The solution is well established with the Mortensen filter in a continuous-time setting, but we seek here its counterpart in a discrete-time context. We demonstrate that it is possible to pursue at the discrete-time level an exact dynamic programming strategy and we find an optimal estimator combining a prediction step using the model and a correction step using the data. This optimal estimator reduces to the discrete-time Kalman estimator when the operators are in fact linear. Furthermore, the strategy that consists of discretizing the least square criterion and then finding the exact estimator at the discrete level allows to determine a new time-scheme for the Mortensen filter which is proven to be consistent and unconditionally stable, with also a consistent and stable discretization of the underlying Hamilton-Jacobi-Bellman equation. This work has resulted in the publication [30].

7.3.2. An iterative method for identifying a stress-free state in image-based biomechanics

Participant: Martin Genet [correspondant].

Continued advances in computational power and methods have enabled image-based biomechanical modeling to become an important tool in basic science, diagnostic and therapeutic medicine, and medical device design. One of the many challenges of this approach, however, is identification of a stress-free reference configuration based on in vivo images of loaded and often prestrained or residually stressed soft tissues and organs. Fortunately, iterative methods have been proposed to solve this inverse problem, among them Sellier's method. This method is particularly appealing because it is easy to implement, converges reasonably fast, and can be coupled to nearly any finite element package. By means of several practical examples, however, we demonstrate that in its original formulation Sellier's method is not optimally fast and may not converge for problems with large deformations. Nevertheless, we can also show that a simple, inexpensive augmentation of Sellier's method based on Aitken's delta-squared process can not only ensure convergence but also significantly accelerate the method. This work has resulted in the publication [31].

7.3.3. A continuum finite strain formulation for finite element image correlation

Participant: Martin Genet [correspondant].

We propose a novel continuum finite strain formulation of the equilibrium gap principle – originally introduced in [Claire, Hild and Roux, 2004, Int. J. Num. Meth. Eng.] at the discrete level for linearized elasticity – used as a regularizer for finite element-based image correlation problems. Consistent linearization and finite element discretization is provided. The method is implemented using FEniCS & VTK, in a freely available Python library. The equilibrium gap constraint regularizes the image correlation problem, even in the presence of noise, and without affecting strain measurements.

7.3.4. Front shape similarity measure for Eikonal PDE data assimilation

Participants: Annabelle Collin [Monc], Philippe Moireau [correspondant].

We present a shape-oriented data assimilation strategy suitable for front-tracking problems through the example of wildfire. The concept of "front" is used to model, at regional scales, the burning area delimitation that moves and undergoes shape and topological changes under heterogeneous orography, biomass fuel and micrometeorology. The simulation-observation discrepancies are represented using a front shape similarity measure inspired from image processing and based on the Chan-Vese contour fitting functional. We show that consistent corrections of the front location and uncertain physical parameters can be obtained using this measure applied on a level-set fire growth model solving for an eikonal equation. This study involves a Luenberger observer for state estimation, including a topological gradient term to track multiple fronts, and a reduced-order Kalman filter for joint parameter estimation. We also highlight the need – prior to parameter estimation – for sensitivity analysis based on the same discrepancy measure, and for instance using polynomial chaos metamodels, to ensure that a meaningful inverse solution is achieved. The performance of the shape-oriented data assimilation strategy is assessed on a synthetic configuration subject to uncertainties in front initial position, near-surface wind magnitude and direction. The use of a robust front shape similarity measure paves the way toward the direct assimilation of infrared images and is a valuable asset in the perspective of data-driven wildfire modeling. This work has resulted in the publication [32].

7.3.5. The mechanism of monomer transfer between two distinct PrP oligomers

Participants: Aurora Armiento, Marie Doumic [Mamba], Philippe Moireau [correspondant].

In mammals, Prion pathology refers to a class of infectious neuropathologies whose mechanism is based on the self-perpetuation of structural information stored in the pathological conformer. The characterisation of the PrP folding landscape has revealed the existence of a plethora of pathways conducing to the formation of structurally different assemblies with different biological properties. However, the biochemical interconnection between these diverse assemblies remains unclear. The PrP oligomerisation process leads to the formation of neurotoxic and soluble assemblies called O1 oligomers with a high size heterodispersity. By combining the measurements in time of size distribution and average size with kinetic models and data assimilation, we revealed the existence of at least two structurally distinct sets of assemblies, termed Oa and Ob, forming O1 assemblies. We propose a kinetic model representing the main processes in prion aggregation pathway: polymerisation, depolymerisation, and disintegration. The two groups interact by exchanging monomers through a disintegration process that increases the size of Oa. Our observations suggest that PrP oligomers constitute a highly dynamic population. This work has resulted in the publication [14].

7.3.6. Joint-state and parameters estimation using ROUKF for HIV mechanistic models

Participants: Annabelle Collin [Monc], Philippe Moireau [correspondant], Mélanie Prague [Sism].

Various methods have been used in the statistical field to estimate parameters in mechanistic models. In particular, an approach based on penalised likelihood for the estimation of parameters in ordinary differential equations with non linear models on parameters (ODE-NLME) has proven successful. For instance, we consider the NIMROD program as a benchmark for estimation in these models. However, such an approach is time consuming. To circumvent this problem, we consider data assimilation approaches that historically arose in the context of geophysics. Here, we propose a Luenberger (also called nudging) state observer coupled with a parameter Kalman-based observer (RoUKF filter, also called SEIK filter) to perform a joint state and parameter estimation on a dataset composed of longitudinal observations of biomarkers for multiple patients. We compare these methods in terms of performances and computation time. We discuss how the concept of random effect can be modeled using Kalman-based filter and its limitations. We illustrate both methods in simulation and on two datasets (the ALBI ANRS 070 trial and the Aquitaine cohort observational data) using an HIV mechanistic model.

7.4. Experimental Assessments

7.4.1. Microstructural interpretation of mouse skin mechanics from multiscale characterization

Participant: Jean-Marc Allain [correspondant].

Skin is a complex, multi-layered organ, with important functions in the protection of the body. The dermis provides structural support to the epidermal barrier, and thus has attracted a large number of mechanical studies. As the dermis is made of a mixture of stiff fibres embedded in a soft non-fibrillar matrix, it is classically considered that its mechanical response is based on an initial alignment of the fibres, followed by the stretching of the aligned fibres. Using a recently developed set-up combining multiphoton microscopy with mechanical assay, we imaged the fibres network evolution during dermis stretching. These observations, combined with a wide set of mechanical tests, allowed us to challenge the classical microstructural interpretation of the mechanical properties of the dermis: we observed a continuous alignment of the collagen fibres along the stretching. All our results can be explained if each fibre contributes by a given stress to the global response. This plastic response is likely due to inner sliding inside each fibre. The non-linear mechanical response is due to structural effects of the fibres network in interaction with the surrounding non-linear matrix. This multiscale interpretation explains our results on genetically-modified mice with a simple alteration of the dermis microstructure. This work has resulted in the publication [27].

7.4.2. Affine kinematics in planar fibrous connective tissues: an experimental investigation

Participants: Jean-Sébastien Affagard, Jean-Marc Allain [correspondant].

The affine transformation hypothesis is usually adopted in order to link the tissue scale with the fibers scale in structural constitutive models of fibrous tissues. Thanks to the recent advances in imaging techniques, such as multiphoton microscopy, the microstructural behavior and kinematics of fibrous tissues can now be monitored at different stretching within the same sample. Therefore, the validity of the affine hypothesis can be investigated. In this study, the fiber reorientation predicted by the affine assumption is compared with experimental data obtained during mechanical tests on skin and liver capsule coupled with microstructural imaging using multiphoton microscopy. The values of local strains and the collagen fibers orientation measured at increasing loading levels are used to compute a theoretical estimation of the affine reorientation of collagen fibers. The experimentally measured reorientation of collagen fibers during loading could not be successfully reproduced with this simple affine model. It suggests that other phenomena occur in the stretching process of planar fibrous connective tissues, which should be included in structural constitutive modeling approaches. This work has resulted in the publication [22].

7.4.3. Improving the experimental protocol for the identification of skin mechanical behavior

Participants: Jean-Sébastien Affagard, Florent Wijanto, Jean-Marc Allain [correspondant].

Mechanical properties of the skin, the external organ of the human body, are important for many applications such as surgery or cosmetics. Due to the highly hierarchical structure of the tissue, it is interesting to develop microstructural models that have better predictability and should reduce the consequences of sample variability. However, these models generally include a quite large number of mechanical parameters. Therefore, complex assays are required to achieve a proper identification of the microstructural models. We investigated in this study the best experimental protocol to identify a nonlinear, anisotropic, model of skin behavior, namely, the Holzapfel law, using displacement field and force measurements. This was done through a sensitivity analysis of the different parameters. We determined first the optimal assay, which appears to be a biaxial test with an alternated loading: first a stretch in one direction, then in the perpendicular one, and so on. To further improve the quality of the assay, we also determined the optimal geometry. Interestingly, slightly asymmetric geometries are more adequate than symmetric ones, while being easier to realise. This work has resulted in the publication [13].

7.4.4. How aging impacts skin biomechanics: a multiscale study in mice

Participants: Jean-Sébastien Affagard, Jean-Marc Allain [correspondant].

Skin aging is a complex process that strongly affects the mechanical behavior of skin. This study aims at deciphering the relationship between age-related changes in dermis mechanical behavior and the underlying changes in dermis microstructure. To that end, we use multiphoton microscopy to monitor the reorganization of dermal collagen during mechanical traction assays in ex vivo skin from young and old mice. The simultaneous variations of a full set of mechanical and microstructural parameters are analyzed in the framework of a multiscale mechanical interpretation. They show consistent results for wild-type mice as well as for genetically-modified mice with modified collagen V synthesis. We mainly observe an increase of the tangent modulus and a lengthening of the heel region in old murine skin from all strains, which is attributed to two different origins that may act together: (i) increased cross-linking of collagen fibers and (ii) loss of water due to proteoglycans deterioration, which impedes inner sliding within these fibers. In contrast, the microstructure reorganization upon stretching shows no age-related difference, which can be attributed to opposite effects of the decrease of collagen content and of the increase of collagen cross-linking in old mice. This work has resulted in the publication [28].

7.4.5. Recent advances in studying single bacteria and biofilm mechanics

Participant: Jean-Marc Allain [correspondant].

Bacterial biofilms correspond to surface-associated bacterial communities embedded in hydrogel-like matrix, in which high cell density, reduced diffusion and physico-chemical heterogeneity play a protective role and induce novel behaviors. We made a summary of the recent advances on the understanding of how bacterial mechanical properties, from single cell to high-cell density community, determine biofilm three-dimensional growth and eventual dispersion, and we attempt to draw a parallel between these properties and the mechanical properties of other well-studied hydrogels and living systems. This work has resulted in the publication [18].

7.5. Clinical Applications

7.5.1. Assessment of atrioventricular valve regurgitation using cardiac modeling

Participants: Radomir Chabiniok [correspondant], Philippe Moireau, Dominique Chapelle.

In this work, we introduce the modeling of atrioventricular valve regurgitation in a spatially reduced-order biomechanical heart model. The model can be fast calibrated using non-invasive data of cardiac magnetic resonance imaging and provides an objective measure of contractile properties of the myocardium in the volume overloaded ventricle, for which the real systolic function may be masked by the significant level of the atrioventricular valve regurgitation. After demonstrating such diagnostic capabilities, we show the potential of modeling to address some clinical questions concerning possible therapeutic interventions for specific patients. The fast running of the model allows targeting specific questions of referring clinicians in a clinically acceptable time. The work was presented at the "Functional Imaging and Modeling of the Heart" conference (FIMH 2017, Toronto, Canada) and is included in the conference proceedings [35].

7.5.2. Model for the dobutamine response in exercise-induced failure of the Fontan circulation

Participants: Radomir Chabiniok [correspondant], Philippe Moireau, Dominique Chapelle.

Understanding physiological phenomena and mechanisms of failure in congenital heart diseases is often challenging due to the complex hemodynamics and high inter-patient variations in anatomy and function. Computational modeling techniques have the potential to greatly improve the understanding of these complex diseases and provide patient-specific clues on mechanisms of deterioration and impact of treatments. This work employs a reduced 0D biomechanical heart model coupled with venous return to capture various key pathophysiological phenomena observed in patients with completed Fontan circulation – a complex surgically established circulation used to palliate patients in whom only one of the two ventricles is functionally able to support the vascular system – with exercise-induced heart failure during dobutamine stress. The framework we propose is fast, efficient and well-suited to the type of pathology and available clinical data obtained by a combined cardiac catheterization and magnetic resonance imaging exam. We demonstrate that the outcomes of modeling are a valuable addition to the current clinical diagnostic investigations and explain patient-specific exercise hemodynamics, identify potential mechanisms of Fontan failure, and enable evaluation of a potential new therapy – selective heart rate modulation – in the treatment of patients with Fontan circulation. The paper is currently in preparation.

7.5.3. Heart and vessels modeling with data interaction for monitoring anesthetized patients

Participants: Arthur Le Gall, Radomir Chabiniok [correspondant], Fabrice Vallée, Dominique Chapelle.

By using mathematical models of heart and vessels developed in the team, we aim at improving intra-operative cardio-vascular safety of anesthetized patients. The patient-specific models, calibrated by echocardiography images and fed by continuous monitoring of aortic arterial pressure and aortic cardiac outflow would allow us to: 1) diagnose pathophysiological modifications associated with changes in the cardio-vascular state; 2) predict the drug response of the patient before the administration of the vaso-active treatment.

7.5.4. Intra-operative monitoring of cardiac afterload

Participants: Arthur Le Gall, Fabrice Vallée [correspondant].

General anesthesia leads to alterations of the cardiovascular system. Intra-operative arterial hypotension is linked to post-operative complications, but using vasopressors to treat arterial hypotension has shown conflicting results. Vasopressors act mainly by elevating cardiac afterload, which could be deleterious in fragile patients, in case of excessive response. Moreover, differences among the most used vasopressors have been observed in vivo [34]. The choice of vasopressor could be important to improve our patients' care. Consequently, we proposed a tool (Velocity-Pressure Loops) to continuously quantify changes in cardiac afterload [33]. Although the first work involves invasive measurement of aortic blood pressure and cardiac outflow, consistent results have been observed when Velocity-Pressure Loops were obtained by a radial arterial catheter with a mathematical transform function [23]. Those findings allow the usage of the Velocity-Pressure Loop without addition of any invasive device.

7.5.5. Review on extra-corporeal circulation

Participant: Arthur Le Gall [correspondant].

This clinical review [26] aims at describing the issues of the management of extra-corporeal membrane oxygenation (ECMO) in the Intensive Care Unit (ICU). From pathophysiology to the description of the impact on mortality, this document shows a global picture of current clinical practices.

7.5.6. *On the importance of consistency in cardiac timings measurements*

Participants: Arthur Le Gall, Alexandre Laurin, Fabrice Vallee [correspondant].

With the contribution of Denis Chemla, professor of Cardiology at Bicêtre Hospital, we presented this work at the CinC conference in Rennes [36]. In this work, we emphasize the need for a consistent method to measure systolic period duration, which is related to cardiac afterload and could be used to quantify arterial pressure amplification phenomenon.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Technical contract with CEA-LIST on coupling strategies between subdomains for transient elasto-dynamics (8keuros)

Contract with the Sensome startup. Aims: feasibility of the measurement of blood clots mechanical properties. (1.6keuros)

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. *ANR*

ANR METIS(ANR-13-BS09-0004-02). Title: “Mechanics of Tissues: multiscale structural approach of Ehlers-Danlos Syndrome”. Involved research groups: LMS (Ecole Polytechnique, CNRS, Mines ParisTech, PI: Jean-Marc ALLAIN), LOB - Optics and Biosciences Laboratory (Ecole Polytechnique, CNRS, INSERM), IGFL - Institut de Génétique Fonctionelle de Lyon (ENS Lyon, Université Lyon 1, CNRS, INRA). Total amount of the grant: 200k€ for the team. The METIS project is dedicated to the study of the biomechanics of connective tissues. Soft connective tissues such as skin, tendon or cornea are made of more than 90% of extracellular matrix proteins, fibrillar collagens being by far the predominant component. The rationale of this project is to understand the link between the microstructure of connective tissues and their macroscopic mechanical properties. To achieve this, observations of the fibrilar collagen will be done at different levels of stretch, while recording the mechanical properties. The consequences of change in the microstructure will also be explored through mutants mimicking the Ehler-Danlos syndrome, but also aging or wound-healing experiments. The project was completed on September 30th 2017 (4 years project).

9.1.2. *Other funding*

IPM-MS project (for Imagerie Polarimétrique de Mueller pour la réalisation d’un système original de caractérisation des propriétés mécaniques des Matériaux Structurés). 50k€ funded by the LABEX Lasips. This project, which involves the LPICM laboratory (Ecole Polytechnique, CNRS), the LMS (Ecole Polytechnique, CNRS, Mines ParisTech) and the Centre des Matériaux (Mines ParisTech), aims at developing an optical tool to study the link between the mechanical properties of a material and its hierarchical organization. Despite the development of new methods to observe the microstructure, one of the limitations is the number of observations that can be obtained on a given sample in a realistic experimental time. To overcome this difficulty, we are planning to use the Mueller polarimetry to obtain at a fast rate (a few frames per second, compared to a few frames per half-hour) relevant information on the local anisotropy of biological (heart, skin) and composite (short fibers composite) samples.

G. Bureau, software engineer in the team, is funded by an Inria Reo industrial contract with Kephalios, a startup working on innovative artificial valves devices.

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

9.2.1.1. VP2HF

Title: Computer model derived indices for optimal patient-specific treatment selection and planning in Heart Failure

Programm: FP7

Duration: October 2013 - March 2017

Coordinator: King's College London (UK)

Inria contact: Dominique Chapelle

Abstract: Heart failure (HF) is one of the major health issues in Europe affecting 6 million patients and growing substantially because of the aging population and improving survival following myocardial infarction. The poor short to medium term prognosis of these patients means that treatments such as cardiac re-synchronisation therapy and mitral valve repair can have substantial impact. However, these therapies are ineffective in up to 50% of the treated patients and involve significant morbidity and substantial cost. The primary aim of VP2HF is to bring together image and data processing tools with statistical and integrated biophysical models mainly developed in previous VPH projects, into a single clinical workflow to improve therapy selection and treatment optimisation in HF.

9.3. International Initiatives

9.3.1. Inria International Partners

9.3.1.1. Informal International Partners

We have started a collaboration with the University of Texas Southwestern Medical Center in Dallas. A joint PhD student based at Inria and funded by UTSW is starting in October 2017. An associated team proposal has been submitted in October 2017.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

9.4.1.1. PhD exchange program

J. Albella, PhD student at University of Santiago de Compostela, has spent 3 months in M3DISIM, working with S. Imperiale on numerical methods for elastodynamics wave propagation.

E. Bertoberoglu, PhD Student at ETH Zurich, has spent multiple weeks in M3DISIM to work with M. Genet on computational models of growth and remodeling of the heart, validated on MRI data acquired at ETH Zurich.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

M. Genet, Session chair on Biomechanics at CSMA2017

M. Genet, Session chair on Software and Algorithms at FEniCS-2017

10.1.1.2. Member of the Organizing Committees

D. Chapelle, Member of the Paris-Saclay Biomechanical Seminar organizing committee

M. Genet, Member of the CSMA-2017 organizing committee

P. Le Tallec, Vice-president of the CSMA2017 (bi-annual national conference on computational structural mechanics) organizing committee

P. Moireau, Member of the organizing committee of the Inria-Saclay teams (Poems-M3disim-Defi) scientific computing seminar

10.1.2. Scientific Events Selection

10.1.2.1. Member of the Conference Program Committees

D. Chapelle, Member of the Conference Program Committee of FIMH 2017

M. Genet, Member of the Conference Program Committee of CSMA-2017

P. Le Tallec, Member of the Conference Program Committee of CSMA-2017

10.1.2.2. Reviewer

D. Chapelle, reviewer for FIMH 2017

M. Genet, reviewer for CSMA-2017

P. Le Tallec, reviewer for CSMA-2017

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

D. Chapelle, Member of the editorial board of journal *Computers & Structures*

D. Chapelle, Member of the editorial board of journal *ESAIM: M2AN*

10.1.3.2. Reviewer - Reviewing Activities

J.-M. Allain, reviewer for “Journal of the Mechanical Behavior of Biomedical Materials” and “Soft Matter journal”

R. Chabiniok, reviewer for “Journal of Computer Methods and Programs in Biomedicine”, “Computers and Structures” and “Journal of Cardiovascular Translational Research”

D. Chapelle, reviewer for “Biomechanics and Modeling in Mechanobiology”, “Computer Methods in Applied Mechanics and Engineering”, “Computers & Structures”

M. Genet, reviewer for “Acta Biomaterialia”

S. Imperiale, reviewer for “Journal of Computers and mathematics with application”, “Journal of Numerical Methods for Partial Differential Equations” “Journal of Methods and Algorithms for Scientific Computing section”, “Journal of Geophysics” and “Journal of Computational Physics”

P. Moireau, reviewer for “ANR” and “Bulletin of Mathematical Biology”

10.1.4. Invited Talks

D. Chapelle, invited seminars at Centrale-Supelec (Jan. 19th) and WIAS Berlin (Sept. 18th)

P. Moireau, “A discrete-time optimal filtering approach for nonlinear systems as a stable discretization of the Mortensen observer” , Workshop on Optimal Control of Dynamical Systems

10.1.5. Leadership within the Scientific Community

J.-M. Allain, Member of Society of Experimental Mechanics and of Biophysical Society

J.-M. Allain, Member of the Academic Council of Université Paris-Saclay, France

D. Chapelle, Head of Science of Inria Saclay-Ile-de-France, and member of the Inria Evaluation Committee

D. Chapelle, Member of the board of directors of the VPH Institute, and of the Avicenna Alliance

D. Chapelle, Member of the steering committee of the BioMedical Engineering Institute coordinated by Ecole Polytechnique

A. Le Gall, Chair of youth committee of SFAR (French Society of Anesthesia and Reanimation)

P. Le Tallec, Director of the LMS (Solid Mechanics Laboratory) Ecole Polytechnique

P. Le Tallec, President of the Mechanics department at University Paris Saclay

P. Moireau, Member of the steering committee of Department of Mathematics of Université Paris Saclay and Jacques Hadamard Foundation

10.1.6. Scientific Expertise

R. Chabiniok, Honorary medical consultant at Saint-Thomas hospital (King's College London)

S. Imperiale, Consultant for CEA (The French Alternative Energies and Atomic Energy Commission)

P. Le Tallec, Consultant for CEA

P. Le Tallec, Consultant for Michelin

P. Moireau, Reviewer for ANR

10.1.7. Research Administration

J.-M. Allain, Responsibility of the teaching experimental center (mechanics), at École Polytechnique

J.-M. Allain, Scientific Advisory Board, chair BioMecAM at ENSAM

J.-M. Allain, Responsible of the "Mechanics and Material for the bio" at the Fédération Francilienne de Mécanique

D. Chapelle, Head of Science of Inria Saclay-Ile-de-France, and member of the Inria Evaluation Committee

P. Le Tallec, Director of the LMS (Solid Mechanics Laboratory) Ecole Polytechnique

P. Le Tallec, in charge of 3rd year program in Mechanics and Energy at Ecole Polytechnique

P. Le Tallec, in charge of the André Citroën Chair

P. Le Tallec, in charge of the Mechanics department at University Paris Saclay

P. Le Tallec, President of the academic senate of the Ecole Polytechnique

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Bachelor: J.-M. Allain, "Introduction projects to physics", 15h, (L1), École Polytechnique, France

Bachelor: F. Caforio,"PEIP1 S2 M2 – Mathematical analysis in two and three dimensions, linear algebra in Rn", 22h, (L1), Polytech Paris-Sud, France

Bachelor: F. Caforio,"Math 255 – Differential calculus for physics (mathematical analysis in two and three dimensions)", 42h, (L2), Université Paris-Sud, France

Bachelor: M. Genet, "MEC435 – Modélisation et Simulation en Mécanique Industrielle", 32h, (L3), École Polytechnique, France

Bachelor: M. Genet,"MEC430 – Mécanique des Milieux Continus I", 32h, (L3), École Polytechnique, France

Bachelor: M. Genet, "MEC431 – Mécanique des Milieux Continus II", 32h, (L3), École Polytechnique, France

Bachelor: S. Imperiale, "MA102 – Analyse pour les EDP", 24h, (L3), ENSTA ParisTech, France

Bachelor: S. Imperiale, "MA104 – Analyse complexe", 12h, (L3), ENSTA ParisTech, France

Bachelor: F. Kimmig, “Mathematical methods for physics”, 32h, (L1) level, École Polytechnique, France

Bachelor: F. Kimmig, “Introduction projects to physics”, 6h, (L1), École Polytechnique, France

Bachelor: F. Kimmig, “Modeling and simulation in industrial mechanics”, 32h, (L3), École Polytechnique, France

Bachelor: P. Le Tallec, “MEC 431 – Mécanique des Milieux Continus 2”, (L3), École Polytechnique, France

Bachelor: P. Le Tallec, “Continuum mechanics”, 32h, (L3), Shanghai ParisTech, China

Bachelor: P. Moireau, “MAP 411 – Approximation numérique et optimisation”, 32 h, (L3), Ecole Polytechnique, France

Master: P. Moireau, “MA103 – Introduction aux EDP et à la méthode des différences finies”, 14h, (L3), ENSTA ParisTech, France

Bachelor: F. Wijanto, “Modeling and simulation in industrial mechanics” 32 h, (L3), École Polytechnique, France

Master: J.-M. Allain, “conputational fluid dynamics”, 36h, (M1), Ecole Polytechnique, France

Master: J.-M. Allain, “Cellular Motility”, 32h, (M2), Ecole Polytechnique, France

Master: D. Chapelle: “Biomechanical Modeling of Active Tissues”, 33h, (M2), Université Paris-Saclay, France

Master: S. Imperiale, “MA2610 Calcul Scientifique – Mécanique des solides”, 6h, (M1), Central/Supelec, France

Master: S. Imperiale, “Simnum – Programmation C++”, 18h, (M1), ENSTA ParisTech, France

Master: P. Le Tallec, “M4S – Numerical methods in non linear mechanics”,27h, (M2), École Polytechnique, France

Master: P. Le Tallec, “Nuclear Energy on Continuum Mechanics”,15h, (M2), INSTN, France

Master: P. Moireau, “MAP-Ann1 – La méthode des éléments finis”, 21h, (M1), ENSTA ParisTech, France

Master: P. Moireau, “Biomechanical Modeling of Active Tissues”, 12h, (M2), Université Paris-Saclay, France

Master: P. Moireau, “Méthodes et problèmes inverses en dynamique des populations”, 24h, (M2), UPMC, France

10.2.2. Supervision

PhD: Aurora Armiento, “Inverse problems and data assimilation methods applied to protein depolymerisation”, supervisors: M. Doumic and P. Moireau, Paris-Descartes University, defended Jan. 13th

PhD in progress: Ustim Khristenko, “Méthodes mathématiques et numériques pour la modélisation et le calcul des états établis cycliques en mécanique non-linéaire”, started Oct. 2014, supervised: P. Le Tallec

PhD in progress: Florent Wijanto, “Modélisation multi-échelle des fibres de collagènes”, started: Sept. 2015, supervisors: J.-M. Allain and M. Caruel

PhD in progress: Federica Caforio, “Modélisation mathématique et numérique de la propagation d’ondes élastique dans le coeur”, started: Nov. 2015, supervisors: D. Chapelle and S. Imperiale

PhD in progress: François Kimmig, “Multi-scale modeling of muscle contraction – From stochastic dynamics of molecular motors to continuum mechanics, in interaction with experimental assays”, started: Sept. 2016, supervisors: M. Caruel and D. Chapelle

PhD in progress: Arthur Le Gall, “Application of biomechanical heart modeling in hemodynamic monitoring of increased risk patients during anesthesia using clinical data”, started Nov. 2016, supervisors: D. Chapelle, E. Gayat (UP7 Paris Diderot) and R. Chabiniok

PhD in progress: Ezgi Berberoglu, “Image-guided Computational Cardiac Mechanics”, started: Jan. 2016 supervisors: Martin Genet and Sebastian Kozerke (ETH, Zurich, Switzerland)

PhD in progress: Cécile Patte, “Modélisation poro-mécanique multi-échelle du poumon, de la respiration au remodelage chronique associé à la fibrose pulmonaire”, started: Nov. 2017, supervisors: M. Genet and D. Chapelle

PhD in progress: Marija Gusseva, “Cardiac Biomechanical Modeling of Chronic Right Ventricular Loading”, started Dec. 2017, supervised: R. Chabiniok, D. Chapelle, T. Hussain (UTSW Medical Center Dallas)

10.2.3. Juries

D. Chapelle, Habilitation Jury of Marcela Szopos, Strasbourg University, President, Dec. 1st

P. Moireau, PhD Jury of Aurora Armiento, Paris-Descartes University, PhD Advisor, Jan. 13th

P. Moireau, PhD Jury of Chloé Audebert, Paris-6 University, Reviewer, Febr. 24th

P. Moireau, PhD Jury of Pauline Bernard, Paris-6 University, President, Nov. 20th

P. Moireau, PhD Jury of Roch Mollero, University of Nice-Sophia Antipolis, Dec. 19th

10.3. Popularization

J.-M. Allain, Jeudi de la Recherche de l’X: popularisation talk on research with medical application. June 2017

J.-M. Allain, 3 articles in popularisation journals (quotidien du médecin, le concours médical, biotech.info) by journalists

D. Chapelle, interview for “Usine Nouvelle”, Dec. 12th

A. Laurin, Inria Saclay inauguration demo and associated news article in “Le Parisien”, Jan. 2017

P. Moireau, Créteil district “Olympiades de Mathématiques” award ceremony, May 2017

R. Chabiniok, Review paper promoting Cardiovascular MRI particularly of congenital heart diseases for Czech and Slovak physicians [17]

A. Laurin, P. Moireau, D. Chapelle, demo at 50 ans de l’Inria, Nov. 7th and 8th

11. Bibliography

Major publications by the team in recent years

[1] R. Chabiniok, P. Moireau, P.-F. Lesault, A. Rahmouni, J.-F. Deux, D. Chapelle. *Estimation of tissue contractility from cardiac cine-MRI using a biomechanical heart model*, in "Biomechanics and Modeling in Mechanobiology", 2012, vol. 11, n° 5, p. 609-630 [*DOI :* 10.1007/s10237-011-0337-8], http://hal.inria.fr/hal-00654541.

[2] D. Chapelle, K. Bathe. *The Finite Element Analysis of Shells - Fundamentals - Second Edition*, Computational Fluid and Solid Mechanics, Springer, 2011, 410 [*DOI :* 10.1007/978-3-642-16408-8], http://hal.inria.fr/hal-00654533.

[3] D. Chapelle, N. Cîndea, P. Moireau. *Improving convergence in numerical analysis using observers - The wave-like equation case*, in "Mathematical Models and Methods in Applied Sciences", 2012, vol. 22, n° 12 [*DOI :* 10.1142/S0218202512500406], http://hal.inria.fr/inria-00621052.

[4] D. CHAPELLE, P. LE TALLEC, P. MOIREAU, M. SORINE.*An energy-preserving muscle tissue model: formulation and compatible discretizations*, in "International Journal for Multiscale Computational Engineering", 2012, vol. 10, n^{o} 2, p. 189-211 [*DOI :* 10.1615/INTJMULTCOMPENG.2011002360], http://hal.inria.fr/hal-00678772.

[5] D. CHAPELLE, P. MOIREAU.*General coupling of porous flows and hyperelastic formulations – From thermodynamics principles to energy balance and compatible time schemes*, in "European Journal of Mechanics - B/Fluids", 2014, vol. 46, p. 82-96, Updated version of previously published research report [*DOI :* 10.1016/J.EUROMECHFLU.2014.02.009], https://hal.inria.fr/inria-00520612.

[6] B. LYNCH, S. BANCELIN, C. BONOD-BIDAUD, J.-B. GUEUSQUIN, F. RUGGIERO, M.-C. SCHANNE-KLEIN, J.-M. ALLAIN.*A novel microstructural interpretation for the biomechanics of mouse skin derived from multiscale characterization*, in "Acta Biomaterialia", 2017, vol. 50, p. 302-311 [*DOI :* 10.1016/J.ACTBIO.2016.12.051], https://hal.archives-ouvertes.fr/hal-01531321.

[7] P. MOIREAU, D. CHAPELLE, P. LE TALLEC.*Joint state and parameter estimation for distributed mechanical systems*, in "Computer Methods in Applied Mechanics and Engineering", 2008, vol. 197, n^{o} 6-8, p. 659-677 [*DOI :* 10.1016/J.CMA.2007.08.021], http://hal.archives-ouvertes.fr/hal-00175623.

[8] P. MOIREAU, D. CHAPELLE, P. LE TALLEC.*Filtering for distributed mechanical systems using position measurements: perspectives in medical imaging*, in "Inverse Problems", 2009, vol. 25, n^{o} 3, 035010 [*DOI :* 10.1088/0266-5611/25/3/035010], http://hal.archives-ouvertes.fr/hal-00358914.

[9] P. MOIREAU, D. CHAPELLE.*Reduced-order Unscented Kalman Filtering with application to parameter identification in large-dimensional systems*, in "ESAIM - Control Optimisation and Calculus of Variations", 2010, Published online - See also erratum DOI:10.1051/cocv/2011001 [*DOI :* 10.1051/COCV/2010006], http://hal.inria.fr/inria-00550104.

[10] P. MOIREAU.*A Discrete-time Optimal Filtering Approach for Non-linear Systems as a Stable Discretization of the Mortensen Observer*, in "ESAIM: Control, Optimisation and Calculus of Variations", 2017, https://hal.inria.fr/hal-01671271.

[11] M. SERMESANT, R. CHABINIOK, P. CHINCHAPATNAM, T. MANSI, F. BILLET, P. MOIREAU, J.-M. PEYRAT, K. C. WONG, J. RELAN, K. S. RHODE, M. GINKS, P. LAMBIASE, H. DELINGETTE, M. SORINE, C. A. RINALDI, D. CHAPELLE, R. RAZAVI, N. AYACHE.*Patient-Specific Electromechanical Models of the Heart for Prediction of the Acute Effects of Pacing in CRT: a First Validation*, in "Medical Image Analysis", January 2012, vol. 16, n^{o} 1, p. 201-215 [*DOI :* 10.1016/J.MEDIA.2011.07.003], http://hal.inria.fr/inria-00616191.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[12] A. ARMIENTO.*Inverse problems and data assimilation methods applied to protein polymerisation*, Université Paris 7 - Diderot, January 2017, https://hal.inria.fr/tel-01447286.

Articles in International Peer-Reviewed Journal

[13] J.-S. AFFAGARD, F. WIJANTO, J.-M. ALLAIN.*Improving the experimental protocol for a more accurate identification of a given mechanical behavior in a single assay: application to skin*, in "Strain", August 2017 [*DOI :* 10.1111/STR.12236], https://hal.archives-ouvertes.fr/hal-01570071.

[14] A. ARMIENTO, P. MOIREAU, D. MARTIN, N. LEPEJOVA, M. DOUMIC, H. REZAEI.*The mechanism of monomer transfer between two structurally distinct PrP oligomers*, in "PLoS ONE", July 2017, vol. 12, n^{o} 7 [*DOI :* 10.1371/JOURNAL.PONE.0180538], https://hal.archives-ouvertes.fr/hal-01574346.

[15] L. ASNER, M. HADJICHARALAMBOUS, R. CHABINIOK, D. PERESSUTTI, E. SAMMUT, J. WONG, G. CARR-WHITE, R. RAZAVI, A. D. KING, N. P. SMITH, J. LEE, D. NORDSLETTEN.*Patient-specific modeling for left ventricular mechanics using data-driven boundary energies*, in "Computer Methods in Applied Mechanics and Engineering", 2017, vol. 314, p. 269-295 [*DOI :* 10.1016/J.CMA.2016.08.002], https://hal.archives-ouvertes.fr/hal-01576770.

[16] B. BURTSCHELL, D. CHAPELLE, P. MOIREAU.*Effective and energy-preserving time discretization for a general nonlinear poromechanical formulation*, in "Computers and Structures", 2017, vol. 182, p. 313-324 [*DOI :* 10.1016/J.COMPSTRUC.2016.10.022], https://hal.inria.fr/hal-01395508.

[17] R. CHABINIOK, L. SÚKUPOVÁ, D. KAUTZNEROVÁ, J. TINTĚRA.*Magnetic Resonance Imaging of the Heart and Large Vessels – Survey of methods and new perspectives*, in "Česká radiologie", November 2017, p. 1-28, https://hal.archives-ouvertes.fr/hal-01667798.

[18] C. EVEN, C. MARLIÈRE, J.-M. GHIGO, J.-M. ALLAIN, A. MARCELLAN, E. RASPAUD.*Recent advances in studying single bacteria and biofilm mechanics*, in "Advances in Colloid and Interface Science", 2017, 57 [*DOI :* 10.1016/J.CIS.2017.07.026], https://hal.archives-ouvertes.fr/hal-01566801.

[19] F. FUENTES, B. KEITH, L. DEMKOWICZ, P. LE TALLEC.*Coupled variational formulations of linear elasticity and the DPG methodology*, in "Journal of Computational Physics", November 2017, vol. 348, p. 715 - 731 [*DOI :* 10.1016/J.JCP.2017.07.051], https://hal.archives-ouvertes.fr/hal-01576307.

[20] M. HADJICHARALAMBOUS, L. ASNER, R. CHABINIOK, E. SAMMUT, J. WONG, D. PERESSUTTI, E. KERFOOT, A. D. KING, J. LEE, R. RAZAVI, N. SMITH, G. CARR-WHITE, D. NORDSLETTEN.*Non-invasive Model-Based Assessment of Passive Left-Ventricular Myocardial Stiffness in Healthy Subjects and in Patients with Non-ischemic Dilated Cardiomyopathy*, in "Annals of Biomedical Engineering", March 2017, vol. 45, n^{o} 3, p. 605-618, https://hal.inria.fr/hal-01419936.

[21] J. HOERMANN, C. BERTOGLIO, M. KRONBICHLER, M. PFALLER, R. CHABINIOK, W. WALL.*An adaptive Hybridizable Discontinuous Galerkin approach for cardiac electrophysiology*, in "International Journal for Numerical Methods in Biomedical Engineering", 2017, p. 1 - 20, https://hal.archives-ouvertes.fr/hal-01672964.

[22] C. JAYYOSI, J.-S. AFFAGARD, G. DUCOURTHIAL, C. BONOD-BIDAUD, B. LYNCH, S. BANCELIN, F. RUGGIERO, M.-C. SCHANNE-KLEIN, J.-M. ALLAIN, K. BRUYÈRE-GARNIER, M. CORET.*Affine kinematics in planar fibrous connective tissues: an experimental investigation*, in "Biomechanics and Modeling in Mechanobiology", 2017, p. 1–15 [*DOI :* 10.1007/S10237-017-0899-1], https://hal.archives-ouvertes.fr/hal-01529263.

[23] J. JOACHIM, F. VALLÉE, J. MATEO, A. LE GALL, S. LENCK, S. MILLASSEAU, E. HOUDART, A. MEBAZAA, E. GAYAT.*Velocity–pressure loops for continuous assessment of ventricular afterload: influence

of pressure measurement site, in "Journal of Clinical Monitoring and Computing", November 2017, p. 1-8 [*DOI :* 10.1007/S10877-017-0082-3], https://hal.inria.fr/hal-01666355.

[24] M. LANDAJUELA, M. VIDRASCU, D. CHAPELLE, M. A. FERNÁNDEZ.*Coupling schemes for the FSI forward prediction challenge: comparative study and validation*, in "International Journal for Numerical Methods in Biomedical Engineering", vol. 33, n^o 4, e02813, 2017 [*DOI :* 10.1002/CNM.2813], https://hal.inria.fr/hal-01239931.

[25] A. LAURIN, M. G. LLOYD, T. HACHIYA, M. SAITO, V. E. CLAYDON, A. BLABER.*New indices from microneurography to investigate the arterial baroreflex*, in "Physiological Reports", June 2017, vol. 5, n^o 12 [*DOI :* 10.14814/PHY2.13220], https://hal.inria.fr/hal-01555596.

[26] A. LE GALL, A. FOLLIN, B. CHOLLEY, J. MANTZ, N. AISSAOUI, R. PIRRACCHIO.*Veno-Arterial-ECMO in the Intensive Care Unit:From Technical Aspects to Clinical Practice*, in "Anaesthesia Critical Care & Pain Medicine", October 2017, p. 1-43 [*DOI :* 10.1016/J.ACCPM.2017.08.007], https://hal.inria.fr/hal-01629078.

[27] B. LYNCH, S. BANCELIN, C. BONOD-BIDAUD, J.-B. GUEUSQUIN, F. RUGGIERO, M.-C. SCHANNE-KLEIN, J.-M. ALLAIN.*A novel microstructural interpretation for the biomechanics of mouse skin derived from multiscale characterization*, in "Acta Biomaterialia", 2017, vol. 50, p. 302-311 [*DOI :* 10.1016/J.ACTBIO.2016.12.051], https://hal.archives-ouvertes.fr/hal-01531321.

[28] B. LYNCH, C. BONOD-BIDAUD, G. DUCOURTHIAL, J.-S. AFFAGARD, S. BANCELIN, S. PSILODIMITRAKOPOULOS, F. RUGGIERO, J.-M. ALLAIN, M.-C. SCHANNE-KLEIN.*How aging impacts skin biomechanics: a multiscale study in mice*, in "Scientific Reports", December 2017, vol. 7, n^o 1 [*DOI :* 10.1038/S41598-017-13150-4], https://hal.archives-ouvertes.fr/hal-01667964.

[29] N. MITROU, A. LAURIN, T. DICK, J. INSKIP.*A peak detection method for identifying phase in physiological signals*, in "Biomedical Signal Processing and Control", 2017, vol. 31, p. 452 - 462 [*DOI :* 10.1016/J.BSPC.2016.07.001], https://hal.inria.fr/hal-01380875.

[30] P. MOIREAU.*A Discrete-time Optimal Filtering Approach for Non-linear Systems as a Stable Discretization of the Mortensen Observer*, in "ESAIM: Control, Optimisation and Calculus of Variations", 2017, https://hal.inria.fr/hal-01671271.

[31] M. K. RAUSCH, M. GENET, J. D. HUMPHREY.*An augmented iterative method for identifying a stress-free reference configuration in image-based biomechanical modeling*, in "Journal of Biomechanics", 2017, vol. 58, p. 227 - 231 [*DOI :* 10.1016/J.JBIOMECH.2017.04.021], https://hal.archives-ouvertes.fr/hal-01571436.

[32] M. ROCHOUX, A. COLLIN, C. ZHANG, A. TROUVÉ, D. LUCOR, P. MOIREAU.*Front shape similarity measure for shape-oriented sensitivity analysis and data assimilation for Eikonal equation*, in "ESAIM: Proceedings and Surveys", 2017, p. 1-22, https://hal.inria.fr/hal-01625575.

[33] F. VALLÉE, A. LE GALL, J. JOACHIM, O. PASSOUANT, J. MATEO, A. MARI, S. MILLASSEAU, A. MEBAZAA, E. GAYAT.*Beat-by-beat assessment of cardiac afterload using descending aortic velocity–pressure loop during general anesthesia: a pilot study*, in "Journal of Clinical Monitoring and Computing", January 2017, vol. 113, n^o 5, p. 727 - 735 [*DOI :* 10.1007/S10877-017-9982-5], https://hal.inria.fr/hal-01629079.

[34] F. VALLÉE, O. PASSOUANT, A. LE GALL, J. JOACHIM, J. MATEO, A. MEBAZAA, E. GAYAT.*Norepinephrine reduces arterial compliance less than phenylephrine when treating general anesthesia-induced arterial hypotension*, in "Acta Anaesthesiologica Scandinavica Supplementum", July 2017, vol. 61, n^o^ 6, p. 590 - 600 [*DOI :* 10.1111/AAS.12905], https://hal.inria.fr/hal-01666364.

International Conferences with Proceedings

[35] R. CHABINIOK, P. MOIREAU, C. KIESEWETTER, T. HUSSAIN, R. RAZAVI, D. CHAPELLE.*Assessment of atrioventricular valve regurgitation using biomechanical cardiac modeling*, in "FIMH 2017 - 9th international conference on Functional Imaging and Modeling of the Heart", Toronto, Canada, LNCS - Lecture Notes in Computer Science, Springer, June 2017, vol. 10263, p. 401-411, https://hal.archives-ouvertes.fr/hal-01505834.

[36] A. LE GALL, A. LAURIN, F. VALLÉE, D. CHEMLA.*Comparison of Systolic Period Duration using Aortic Flow or Pressure Based Methods in Anesthetized Patients*, in "Computing in cardioogy", Rennes, France, September 2017, https://hal.inria.fr/hal-01580305.

National Conferences with Proceeding

[37] M. PRAGUE, R. THIÉBAUT, P. MOIREAU, A. COLLIN.*Joint-state and parameters estimation using nudging and SEIK filters for HIV mechanistic models*, in "Journée de la statistique francaise", Avignon, France, June 2017, https://hal.inria.fr/hal-01579068.

Conferences without Proceedings

[38] M. GENET, L. C. LEE, S. KOZERKE.*A continuum finite strain formulation of the equilibrium gap regularizer for finite element image correlation*, in "13e`me Colloque National en Calcul des Structures", Giens, France, May 2017, https://hal.archives-ouvertes.fr/hal-01661810.

[39] M. GENET, L. C. LEE, S. KOZERKE.*Equilibrated Warping: Finite Element Image Correlation with Mechanical Regularization*, in "FEniCS'2017", Luxembourg, Luxembourg, June 2017, https://hal.archives-ouvertes.fr/hal-01661825.

Patents and standards

[40] R. CHABINIOK, D. CHAPELLE, A. LE GALL, P. MOIREAU, F. VALLÉE.*Dispositif cardiaque*, 2017, n^o^ 1758006, https://hal.inria.fr/hal-01580331.

Project-Team MEXICO

Modeling and Exploitation of Interaction and Concurrency

IN COLLABORATION WITH: Laboratoire specification et vérification (LSV)

IN PARTNERSHIP WITH:
CNRS

Ecole normale supérieure de Cachan

RESEARCH CENTER
Saclay - Île-de-France

THEME
Proofs and Verification

Table of contents

Project-Team MEXICO

Creation of the Team: 2009 March 01, updated into Project-Team: 2011 January 01

Keywords:

Computer Science and Digital Science:

A6.4.1. - Deterministic control
A6.4.3. - Observability and Controlability
A7.1. - Algorithms
A7.2. - Logic in Computer Science
A8.1. - Discrete mathematics, combinatorics
A8.2. - Optimization
A8.7. - Graph theory
A8.8. - Network science
A8.9. - Performance evaluation

Other Research Topics and Application Domains:

B1.1.2. - Molecular biology
B1.1.3. - Cellular biology
B1.1.11. - Systems biology
B1.1.12. - Synthetic biology
B6.3.1. - Web
B6.3.2. - Network protocols
B6.3.3. - Network Management
B7.1. - Traffic management
B7.2.1. - Smart vehicles

1. Personnel

Research Scientists

Stefan Haar [Team leader, Inria, Senior Researcher, HDR]
Beatrice Berard [Univ Denis Diderot Paris, Senior Researcher, until Aug 2017]
Laurent Fribourg [CNRS, Senior Researcher, HDR]
Matthias Fuegger [CNRS, Researcher]

Faculty Members

Thomas Chatain [Ecole Normale Supérieure Paris-Saclay, Associate Professor, HDR]
Serge Haddad [Ecole Normale Supérieure Paris-Saclay, Professor, HDR]
Claudine Picaronny [Ecole Normale Supérieure Paris-Saclay, Associate Professor]
Stefan Schwoon [Ecole Normale Supérieure Paris-Saclay, Associate Professor, HDR]

External Collaborator

Benoît Barbot [Univ Paris-Est Marne La Vallée]

PhD Students

Yann Duplouy [Institut de recherche technologique System X]
Juraj Kolcak [Inria, from Mar 2017]
Hugues Mandon [Inria]
Adnane Saoud [CNRS]

Administrative Assistants

Thida Iem [Inria, until Sep 2017]

Emmanuelle Perrot [Inria, since Oct 2017]

2. Overall Objectives

2.1. Scientific Objectives

2.1.1. Introduction

In the increasingly networked world, reliability of applications becomes ever more critical as the number of users of, e.g., communication systems, web services, transportation etc., grows steadily. Management of networked systems, in a very general sense of the term, therefore is a crucial task, but also a difficult one.

MExICo strives to take advantage of distribution by orchestrating cooperation between different agents that observe local subsystems, and interact in a localized fashion.

The need for applying formal methods in the analysis and management of complex systems has long been recognized. It is with much less unanimity that the scientific community embraces methods based on asynchronous and distributed models. Centralized and sequential modeling still prevails.

However, we observe that crucial applications have increasing numbers of users, that networks providing services grow fast both in the number of participants and the physical size and degree of spatial distribution. Moreover, traditional *isolated* and *proprietary* software products for local systems are no longer typical for emerging applications.

In contrast to traditional centralized and sequential machinery for which purely functional specifications are efficient, we have to account for applications being provided from diverse and non-coordinated sources. Their distribution (e.g. over the Web) must change the way we verify and manage them. In particular, one cannot ignore the impact of quantitative features such as delays or failure likelihoods on the functionalities of composite services in distributed systems.

We thus identify three main characteristics of complex distributed systems that constitute research challenges:

- *Concurrency* of behavior;
- *Interaction* of diverse and semi-transparent components; and
- management of *Quantitative* aspects of behavior.

2.1.2. Concurrency

The increasing size and the networked nature of communication systems, controls, distributed services, etc. confront us with an ever higher degree of parallelism between local processes. This field of application for our work includes telecommunication systems and composite web services. The challenge is to provide sound theoretical foundations and efficient algorithms for management of such systems, ranging from controller synthesis and fault diagnosis to integration and adaptation. While these tasks have received considerable attention in the *sequential* setting, managing *non-sequential* behavior requires profound modifications for existing approaches, and often the development of new approaches altogether. We see concurrency in distributed systems as an opportunity rather than a nuisance. Our goal is to *exploit* asynchronicity and distribution as an advantage. Clever use of adequate models, in particular *partial order semantics* (ranging from Mazurkiewicz traces to event structures to MSCs) actually helps in practice. In fact, the partial order vision allows us to make causal precedence relations explicit, and to perform diagnosis and test for the dependency between events. This is a conceptual advantage that interleaving-based approaches cannot match. The two key features of our work will be *(i)* the exploitation of concurrency by using asynchronous models with partial order semantics, and *(ii)* distribution of the agents performing management tasks.

2.1.3. Interaction

Systems and services exhibit non-trivial *interaction* between specialized and heterogeneous components. A coordinated interplay of several components is required; this is challenging since each of them has only a limited, partial view of the system's configuration. We refer to this problem as *distributed synthesis* or *distributed control*. An aggravating factor is that the structure of a component might be semi-transparent, which requires a form of *grey box management*.

2.1.4. Quantitative Features

Besides the logical functionalities of programs, the *quantitative* aspects of component behavior and interaction play an increasingly important role.

- *Real-time* properties cannot be neglected even if time is not an explicit functional issue, since transmission delays, parallelism, etc, can lead to time-outs striking, and thus change even the logical course of processes. Again, this phenomenon arises in telecommunications and web services, but also in transport systems.
- In the same contexts, *probabilities* need to be taken into account, for many diverse reasons such as unpredictable functionalities, or because the outcome of a computation may be governed by race conditions.
- Last but not least, constraints on *cost* cannot be ignored, be it in terms of money or any other limited resource, such as memory space or available CPU time.

2.1.5. Evolution and Perspectives

Since the creation of *MExICo*, the weight of *quantitative* aspects in all parts of our activities has grown, be it in terms of the models considered (weighted automata and logics), be it in transforming verification or diagnosis verdict into probabilistic statements (probabilistic diagnosis, statistical model checking), or within the recently started SystemX cooperation on supervision in multi-modal transport systems. This trend is certain to continue over the next couple of years, along with the growing importance of diagnosis and control issues.

In another development, the theory and use of partial order semantics has gained momentum in the past four years, and we intend to further strengthen our efforts and contacts in this domain to further develop and apply partial-order based deduction methods.

As concerns the study of interaction, our progress has been thus far less in the domain of *distributed* approaches than in the analysis of *system composition*, such as in networks of untimed or timed automata. While continuing this line of study, we also intend to turn more strongly towards distributed *algorithms*, namely in terms of parametrized verification methods.

3. Research Program

3.1. Concurrency

Participants: Thomas Chatain, Stefan Haar, Serge Haddad, Stefan Schwoon.

Concurrency; Semantics; Automatic Control ; Diagnosis ; Verification

Concurrency: Property of systems allowing some interacting processes to be executed in parallel.

Diagnosis: The process of deducing from a partial observation of a system aspects of the internal states or events of that system; in particular, *fault diagnosis* aims at determining whether or not some non-observable fault event has occurred.

Conformance Testing: Feeding dedicated input into an implemented system IS and deducing, from the resulting output of I, whether I respects a formal specification S.

3.1.1. Introduction

It is well known that, whatever the intended form of analysis or control, a *global* view of the system state leads to overwhelming numbers of states and transitions, thus slowing down algorithms that need to explore the state space. Worse yet, it often blurs the mechanics that are at work rather than exhibiting them. Conversely, respecting concurrency relations avoids exhaustive enumeration of interleavings. It allows us to focus on 'essential' properties of non-sequential processes, which are expressible with causal precedence relations. These precedence relations are usually called causal (partial) orders. Concurrency is the explicit absence of such a precedence between actions that do not have to wait for one another. Both causal orders and concurrency are in fact essential elements of a specification. This is especially true when the specification is constructed in a distributed and modular way. Making these ordering relations explicit requires to leave the framework of state/interleaving based semantics. Therefore, we need to develop new dedicated algorithms for tasks such as conformance testing, fault diagnosis, or control for distributed discrete systems. Existing solutions for these problems often rely on centralized sequential models which do not scale up well.

3.1.2. Diagnosis

Participants: Stefan Haar, Serge Haddad, Stefan Schwoon.

Fault Diagnosis for discrete event systems is a crucial task in automatic control. Our focus is on *event oriented* (as opposed to *state oriented*) model-based diagnosis, asking e.g. the following questions:
given a - potentially large - *alarm pattern* formed of observations,

- what are the possible *fault scenarios* in the system that *explain* the pattern ?
- Based on the observations, can we deduce whether or not a certain - invisible - fault has actually occurred ?

Model-based diagnosis starts from a discrete event model of the observed system - or rather, its relevant aspects, such as possible fault propagations, abstracting away other dimensions. From this model, an extraction or unfolding process, guided by the observation, produces recursively the explanation candidates.

In asynchronous partial-order based diagnosis with Petri nets [40], [41], [42], one unfolds the *labelled product* of a Petri net model $\mathcal{N}$ and an observed alarm pattern $\mathcal{A}$, also in Petri net form. We obtain an acyclic net giving partial order representation of the behaviors compatible with the alarm pattern. A recursive online procedure filters out those runs *(configurations)* that explain *exactly*$\mathcal{A}$. The Petri-net based approach generalizes to dynamically evolving topologies, in dynamical systems modeled by graph grammars, see [29]

3.1.2.1. Observability and Diagnosability

Diagnosis algorithms have to operate in contexts with low observability, i.e., in systems where many events are invisible to the supervisor. Checking *observability* and *diagnosability* for the supervised systems is therefore a crucial and non-trivial task in its own right. Analysis of the relational structure of occurrence nets allows us to check whether the system exhibits sufficient visibility to allow diagnosis. Developing efficient methods for both verification of *diagnosability checking* under concurrency, and the *diagnosis* itself for distributed, composite and asynchronous systems, is an important field for *MExICo*.

3.1.2.2. Distribution

Distributed computation of unfoldings allows one to factor the unfolding of the global system into smaller *local* unfoldings, by local supervisors associated with sub-networks and communicating among each other. In [41], [31], elements of a methodology for distributed computation of unfoldings between several supervisors, underwritten by algebraic properties of the category of Petri nets have been developed. Generalizations, in particular to Graph Grammars, are still do be done.

Computing diagnosis in a distributed way is only one aspect of a much vaster topic, that of *distributed diagnosis* (see [38], [44]). In fact, it involves a more abstract and often indirect reasoning to conclude whether or not some given invisible fault has occurred. Combination of local scenarios is in general not sufficient: the global system may have behaviors that do not reveal themselves as faulty (or, dually, non-faulty) on any local supervisor's domain (compare [28], [34]). Rather, the local diagnosers have to join all *information* that is available to them locally, and then deduce collectively further information from the combination of their views. In particular, even the *absence* of fault evidence on all peers may allow to deduce fault occurrence jointly, see [46], [47]. Automatizing such procedures for the supervision and management of distributed and locally monitored asynchronous systems is a long-term goal to which *MExICo* hopes to contribute.

3.1.3. Hybrid Systems

Participants: Laurent Fribourg, Serge Haddad.

Hybrid systems constitute a model for cyber-physical systems which integrates continuous-time dynamics (modes) governed by differential equations, and discrete transitions which switch instantaneously from one mode to another. Thanks to their ease of programming, hybrid systems have been integrated to power electronics systems, and more generally in cyber-physical systems. In order to guarantee that such systems meet their specifications, classical methods consist in finitely abstracting the systems by discretization of the (infinite) state space, and deriving automatically the appropriate mode control from the specification using standard graph techniques. These methods face the well-known problem of "curse of dimensionality", and cannot generally treat systems of dimension exceeding 5 or 6. Thanks to the introduction of original compositional techniques [22] as well as finer estimations of integration errors [21], we are now able to control several case studies of greater dimension. Actually, in the real world, many parameters of hybrid models are not known precisely, and require adjustements to experimental data. We plan to elaborate methods based on parameter estimation and machine learning techniques in order to define formal stability criteria and well-posed learning problems in the framework of hybrid systems with nonlinear dynamics.

3.1.4. Contextual Nets

Participant: Stefan Schwoon.

Assuring the correctness of concurrent systems is notoriously difficult due to the many unforeseeable ways in which the components may interact and the resulting state-space explosion. A well-established approach to alleviate this problem is to model concurrent systems as Petri nets and analyse their unfoldings, essentially an acyclic version of the Petri net whose simpler structure permits easier analysis [39].

However, Petri nets are inadequate to model concurrent read accesses to the same resource. Such situations often arise naturally, for instance in concurrent databases or in asynchronous circuits. The encoding tricks typically used to model these cases in Petri nets make the unfolding technique inefficient. Contextual nets, which explicitly do model concurrent read accesses, address this problem. Their accurate representation of concurrency makes contextual unfoldings up to exponentially smaller in certain situations. An abstract algorithm for contextual unfoldings was first given in [30]. In recent work, we further studied this subject from a theoretical and practical perspective, allowing us to develop concrete, efficient data structures and algorithms and a tool (Cunf) that improves upon existing state of the art. This work led to the PhD thesis of César Rodríguez in 2014 .

Contexutal unfoldings deal well with two sources of state-space explosion: concurrency and shared resources. Recently, we proposed an improved data structure, called *contextual merged processes* (CMP) to deal with a third source of state-space explosion, i.e. sequences of choices. The work on CMP [48] is currently at an abstract level. In the short term, we want to put this work into practice, requiring some theoretical groundwork, as well as programming and experimentation.

Another well-known approach to verifying concurrent systems is *partial-order reduction*, exemplified by the tool SPIN. Although it is known that both partial-order reduction and unfoldings have their respective strengths and weaknesses, we are not aware of any conclusive comparison between the two techniques. Spin comes with a high-level modeling language having an explicit notion of processes, communication channels, and variables.

Indeed, the reduction techniques implemented in Spin exploit the specific properties of these features. On the other side, while there exist highly efficient tools for unfoldings, Petri nets are a relatively general low-level formalism, so these techniques do not exploit properties of higher language features. Our work on contextual unfoldings and CMPs represents a first step to make unfoldings exploit richer models. In the long run, we wish raise the unfolding technique to a suitable high-level modelling language and develop appropriate tool support.

3.2. Management of Quantitative Behavior

Participants: Thomas Chatain, Stefan Haar, Serge Haddad.

3.2.1. Introduction

Besides the logical functionalities of programs, the *quantitative* aspects of component behavior and interaction play an increasingly important role.

- *Real-time* properties cannot be neglected even if time is not an explicit functional issue, since transmission delays, parallelism, etc, can lead to time-outs striking, and thus change even the logical course of processes. Again, this phenomenon arises in telecommunications and web services, but also in transport systems.
- In the same contexts, *probabilities* need to be taken into account, for many diverse reasons such as unpredictable functionalities, or because the outcome of a computation may be governed by race conditions.
- Last but not least, constraints on *cost* cannot be ignored, be it in terms of money or any other limited resource, such as memory space or available CPU time.

Traditional mainframe systems were proprietary and (essentially) localized; therefore, impact of delays, unforeseen failures, etc. could be considered under the control of the system manager. It was therefore natural, in verification and control of systems, to focus on *functional* behavior entirely.

With the increase in size of computing system and the growing degree of compositionality and distribution, quantitative factors enter the stage:

- calling remote services and transmitting data over the web creates *delays*;
- remote or non-proprietary components are not “deterministic”, in the sense that their behavior is uncertain.

Time and *probability* are thus parameters that management of distributed systems must be able to handle; along with both, the *cost* of operations is often subject to restrictions, or its minimization is at least desired. The mathematical treatment of these features in distributed systems is an important challenge, which *MExICo* is addressing; the following describes our activities concerning probabilistic and timed systems. Note that cost optimization is not a current activity but enters the picture in several intended activities.

3.2.2. Probabilistic distributed Systems

Participants: Stefan Haar, Serge Haddad, Claudine Picaronny.

3.2.2.1. Non-sequential probabilistic processes

Practical fault diagnosis requires to select explanations of *maximal likelihood*. For partial-order based diagnosis, this leads therefore to the question what the probability of a given partially ordered execution is. In Benveniste et al. [33], [26], we presented a model of stochastic processes, whose trajectories are partially ordered, based on local branching in Petri net unfoldings; an alternative and complementary model based on Markov fields is developed in [43], which takes a different view on the semantics and overcomes the first model’s restrictions on applicability.

Both approaches abstract away from real time progress and randomize choices in *logical* time. On the other hand, the relative speed - and thus, indirectly, the real-time behavior of the system’s local processes - are crucial factors determining the outcome of probabilistic choices, even if non-determinism is absent from the system.

In another line of research [35] we have studied the likelihood of occurrence of non-sequential runs under random durations in a stochastic Petri net setting. It remains to better understand the properties of the probability measures thus obtained, to relate them with the models in logical time, and exploit them e.g. in *diagnosis*.

3.2.2.2. Distributed Markov Decision Processes

Participant: Serge Haddad.

Distributed systems featuring non-deterministic and probabilistic aspects are usually hard to analyze and, more specifically, to optimize. Furthermore, high complexity theoretical lower bounds have been established for models like partially observed Markovian decision processes and distributed partially observed Markovian decision processes. We believe that these negative results are consequences of the choice of the models rather than the intrinsic complexity of problems to be solved. Thus we plan to introduce new models in which the associated optimization problems can be solved in a more efficient way. More precisely, we start by studying connection protocols weighted by costs and we look for online and offline strategies for optimizing the mean cost to achieve the protocol. We have been cooperating on this subject with the SUMO team at Inria Rennes; in the joint work [27]; there, we strive to synthesize for a given MDP a control so as to guarantee a specific stationary behavior, rather than - as is usually done - so as to maximize some reward.

3.2.3. Large scale probabilistic systems

Addressing large-scale probabilistic systems requires to face state explosion, due to both the discrete part and the probabilistic part of the model. In order to deal with such systems, different approaches have been proposed:

- Restricting the synchronization between the components as in queuing networks allows to express the steady-state distribution of the model by an analytical formula called a product-form [32].
- Some methods that tackle with the combinatory explosion for discrete-event systems can be generalized to stochastic systems using an appropriate theory. For instance symmetry based methods have been generalized to stochastic systems with the help of aggregation theory [37].
- At last simulation, which works as soon as a stochastic operational semantic is defined, has been adapted to perform statistical model checking. Roughly speaking, it consists to produce a confidence interval for the probability that a random path fulfills a formula of some temporal logic [49] .

We want to contribute to these three axes: (1) we are looking for product-forms related to systems where synchronization are more involved (like in Petri nets), see [2]; (2) we want to adapt methods for discrete-event systems that require some theoretical developments in the stochastic framework and, (3) we plan to address some important limitations of statistical model checking like the expressiveness of the associated logic and the handling of rare events.

3.2.4. Real time distributed systems

Nowadays, software systems largely depend on complex timing constraints and usually consist of many interacting local components. Among them, railway crossings, traffic control units, mobile phones, computer servers, and many more safety-critical systems are subject to particular quality standards. It is therefore becoming increasingly important to look at networks of timed systems, which allow real-time systems to operate in a distributed manner.

Timed automata are a well-studied formalism to describe reactive systems that come with timing constraints. For modeling distributed real-time systems, networks of timed automata have been considered, where the local clocks of the processes usually evolve at the same rate [45] [36]. It is, however, not always adequate to assume that distributed components of a system obey a global time. Actually, there is generally no reason to assume that different timed systems in the networks refer to the same time or evolve at the same rate. Any component is rather determined by local influences such as temperature and workload.

3.2.4.1. Implementation of Real-Time Concurrent Systems

Participants: Thomas Chatain, Stefan Haar, Serge Haddad.

This was one of the tasks of the ANR ImpRo.

Formal models for real-time systems, like timed automata and time Petri nets, have been extensively studied and have proved their interest for the verification of real-time systems. On the other hand, the question of using these models as specifications for designing real-time systems raises some difficulties. One of those comes from the fact that the real-time constraints introduce some artifacts and because of them some syntactically correct models have a formal semantics that is clearly unrealistic. One famous situation is the case of Zeno executions, where the formal semantics allows the system to do infinitely many actions in finite time. But there are other problems, and some of them are related to the distributed nature of the system. These are the ones we address here.

One approach to implementability problems is to formalize either syntactical or behavioral requirements about what should be considered as a reasonable model, and reject other models. Another approach is to adapt the formal semantics such that only realistic behaviors are considered.

These techniques are preliminaries for dealing with the problem of implementability of models. Indeed implementing a model may be possible at the cost of some transformation, which make it suitable for the target device. By the way these transformations may be of interest for the designer who can now use high-level features in a model of a system or protocol, and rely on the transformation to make it implementable.

We aim at formalizing and automating translations that preserve both the timed semantics and the concurrent semantics. This effort is crucial for extending concurrency-oriented methods for logical time, in particular for exploiting partial order properties. In fact, validation and management - in a broad sense - of distributed systems is not realistic *in general* without understanding and control of their real-time dependent features; the link between real-time and logical-time behaviors is thus crucial for many aspects of *MExICo*'s work.

4. Application Domains

4.1. Telecommunications

Participants: Stefan Haar, Serge Haddad.

Telecommunications

MExICo's research is motivated by problems of system management in several domains, such as:

- In the domain of service oriented computing, it is often necessary to insert some Web service into an existing orchestrated business process, e.g. to replace another component after failures. This requires to ensure, often actively, conformance to the interaction protocol. One therefore needs to synthesize adaptators for every component in order to steer its interaction with the surrounding processes.
- Still in the domain of telecommunications, the supervision of a network tends to move from out-of-band technology, with a fixed dedicated supervision infrastructure, to in-band supervision where the supervision process uses the supervised network itself. This new setting requires to revisit the existing supervision techniques using control and diagnosis tools.

Currently, we have no active cooperation on thes subjects.

4.2. Biological Systems

Participants: Thomas Chatain, Stefan Haar, Serge Haddad, Stefan Schwoon.

We have begun in 2014 to examine concurrency issues in systems biology, and are currently enlarging the scope of our research's applications in this direction. To see the context, note that in recent years, a considerable shift of biologists' interest can be observed, from the mapping of static genotypes to gene expression, i.e. the processes in which genetic information is used in producing functional products. These processes are far from being uniquely determined by the gene itself, or even jointly with static properties of the environment; rather, regulation occurs throughout the expression processes, with specific mechanisms increasing or decreasing the production of various products, and thus modulating the outcome. These regulations are central in understanding cell fate (how does the cell differenciate ? Do mutations occur ? etc), and progress there hinges on our capacity to analyse, predict, monitor and control complex and variegated processes. We have applied Petri net unfolding techniques for the efficient computation of attractors in a regulatory network; that is, to identify strongly connected reachability components that correspond to stable evolutions, e.g. of a cell that differentiates into a specific functionality (or mutation). This constitutes the starting point of a broader research with Petri net unfolding techniques in regulation. In fact, ‚he use of ordinary Petri nets for capturing regulatory network (RN) dynamics overcomes the limitations of traditional RN models : those impose e.g. Monotonicity properties in the influence that one factor had upon another, i.e. always increasing or always decreasing, and were thus unable to cover all actual behaviours (see [75]). Rather, we follow the more refined model of boolean networks of automata, where the local states of the different factors jointly detemine which state transitions are possible. For these connectors, ordinary PNs constitute a first approximation, improving greatly over the literature but leaving room for improvement in terms of introducing more refined logical connectors. Future work thus involves transcending this class of PN models. Via unfoldings, one has access – provided efficient techniques are available – to all behaviours of the model, rather than over-or under-approximations as previously. This opens the way to efficiently searching in particular for determinants of the cell fate : which attractors are reachable from a given stage, and what are the factors that decide in favor of one or the other attractor, etc. Our current research focusses on *cellular reprogramming*.

4.3. Autonomous Vehicles

The validation of safety properties is a crucial concern for the design of computer guided systems, in particular for automated transport systems Our approach consists in analyzing the interactions of a randomized environment (roads, cross-sections, etc.) with a vehicle controller. This requires to :

- define the relevant case studies;
- extend our tool COSMOS to handle general hybrid systems;
- conduct experimentations and analyze their results.

In [SIA2017], we have shown that this approach scales pretty well but with a controller written in C. The next step will be to combine Simulink models with Petri nets since Simulink is widely used for specifying hybrid systems in industry. In order to so, we need to define an operational semantic for Simulink and to design an elegant way for specifying the interface between nets and Simulink models. Then we will implement the solution in Cosmos.

5. Highlights of the Year

5.1. Highlights of the Year

See the 'New results' section.

6. New Software and Platforms

6.1. COSMOS

FUNCTIONAL DESCRIPTION: COSMOS is a statistical model checker for the Hybrid Automata Stochastic Logic (HASL). HASL employs Linear Hybrid Automata (LHA), a generalization of Deterministic Timed Automata (DTA), to describe accepting execution paths of a Discrete Event Stochastic Process (DESP), a class of stochastic models which includes, but is not limited to, Markov chains. As a result HASL verification turns out to be a unifying framework where sophisticated temporal reasoning is naturally blended with elaborate reward-based analysis. COSMOS takes as input a DESP (described in terms of a Generalized Stochastic Petri Net), an LHA and an expression Z representing the quantity to be estimated. It returns a confidence interval estimation of Z, recently, it has been equipped with functionalities for rare event analysis. COSMOS is written in C++

- Participants: Benoît Barbot, Hilal Djafri, Marie Duflot-Kremer, Paolo Ballarini and Serge Haddad
- Contact: Hilal Djafri
- URL: http://www.lsv.ens-cachan.fr/~barbot/cosmos/

6.2. CosyVerif

FUNCTIONAL DESCRIPTION: CosyVerif is a platform dedicated to the formal specification and verification of dynamic systems. It allows to specify systems using several formalisms (such as automata and Petri nets), and to run verification tools on these models.

- Participants: Alban Linard, Fabrice Kordon, Laure Petrucci and Serge Haddad
- Partners: LIP6 - LSV - LIPN (Laboratoire d'Informatique de l'Université Paris Nord)
- Contact: Serge Haddad
- URL: http://www.cosyverif.org/

6.3. Mole

FUNCTIONAL DESCRIPTION: Mole computes, given a safe Petri net, a finite prefix of its unfolding. It is designed to be compatible with other tools, such as PEP and the Model-Checking Kit, which are using the resulting unfolding for reachability checking and other analyses. The tool Mole arose out of earlier work on Petri nets.

- Participant: Stefan Schwoon
- Contact: Stefan Schwoon
- URL: http://www.lsv.ens-cachan.fr/~schwoon/tools/mole/

7. New Results

7.1. Optimal constructions for active diagnosis

Published in [4].

Diagnosis is the task of detecting fault occurrences in a partially observed system. Depending on the possible observations, a discrete-event system may be diagnosable or not. Active diagnosis aims at controlling the system to render it diagnosable. Past research has proposed solutions for this problem, but their complexity remains to be improved. Here, we solve the decision and synthesis problems for active diagnosability, proving that (1) our procedures are optimal with respect to computational complexity, and (2) the memory required for our diagnoser is minimal. We then study the delay between a fault occurrence and its detection by the diagnoser. We construct a memory-optimal diagnoser whose delay is at most twice the minimal delay, whereas the memory required to achieve optimal delay may be highly greater. We also provide a solution for parametrized active diagnosis, where we automatically construct the most permissive controller respecting a given delay.

7.2. Diagnosability of Repairable Faults

Published in [3].

The diagnosis problem for discrete event systems consists in deciding whether some fault event occurred or not in the system, given partial observations on the run of that system. Diagnosability checks whether a correct diagnosis can be issued in bounded time after a fault, for all faulty runs of that system. This problem appeared two decades ago and numerous facets of it have been explored, mostly for permanent faults. It is known for example that diagnosability of a system can be checked in polynomial time, while the construction of a diagnoser is exponential. The present paper examines the case of transient faults, that can appear and be repaired. Diagnosability in this setting means that the occurrence of a fault should always be detected in bounded time, but also before the fault is repaired, in order to prepare for the detection of the next fault or to take corrective measures while they are needed. Checking this notion of diagnosability is proved to be PSPACE-complete. It is also shown that faults can be reliably counted provided the system is diagnosable for faults and for repairs.

7.3. Diagnostic et contrôle de la dégradation des systèmes probabilistes

Published in [18].

Le diagnostic actif est opéré par un contrôleur en vue de rendre un système diagnosticable. Afin d'éviter que le contrôleur ne dégrade trop fortement le système, on lui affecte généralement un second objectif en termes de qualité de service. Dans le cadre des systèmes probabilistes, une spécification possible consiste à assurer une probabilité positive qu'une exécution infinie soit correcte, ce qu'on appelle le diagnostic actif sûr. Nous introduisons ici deux spécifications alternatives. La gamma-correction du système affecte à une exécution une valeur de correction dépendant d'un facteur de décote gamma et le contrôleur doit assurer une valeur moyenne supérieure à un seuil fixé. La alpha-dégradation requiert qu'asymptotiquement, à chaque unité de temps une proportion supérieure à alpha des exécutions jusqu'alors correctes le demeure. D'un point de vue sémantique, nous explicitons des liens significatifs entre les différentes notions. Algorithmiquement, nous établissons la frontière entre décidabilité et indécidabilité des problèmes et dans le cas positif nous exhibons la complexité précise ainsi qu'une synthèse, potentiellement à mémoire infinie.

7.4. The Complexity of Diagnosability and Opacity Verification for Petri Nets

Published in [7].

Diagnosability and opacity are two well-studied problems in discrete-event systems. We revisit these two problems with respect to expressiveness and complexity issues. We first relate different notions of diagnosability and opacity. We consider in particular fairness issues and extend the definition of Germanos et al. [ACM TECS, 2015] of weakly fair diagnosability for safe Petri nets to general Petri nets and to opacity questions. Second, we provide a global picture of complexity results for the verification of diagnosability and opacity. We show that diagnosability is NL-complete for finite state systems, PSPACE-complete for safe Petri nets (even with fairness), and EXPSPACE-complete for general Petri nets without fairness, while non diagnosability is inter-reducible with reachability when fault events are not weakly fair. Opacity is ESPACE-complete for safe Petri nets (even with fairness) and undecidable for general Petri nets already without fairness.

7.5. Probabilistic Disclosure: Maximisation vs. Minimisation

Published in [8].

We consider opacity questions where an observation function provides to an external attacker a view of the states along executions and secret executions are those visiting some state from a fixed subset. Disclosure occurs when the observer can deduce from a finite observation that the execution is secret, the ε-disclosure variant corresponding to the execution being secret with probability greater than 1 - ε. In a probabilistic and non deterministic setting, where an internal agent can choose between actions, there are two points of view, depending on the status of this agent: the successive choices can either help the attacker trying to disclose the secret, if the system has been corrupted, or they can prevent disclosure as much as possible if these choices are part of the system design. In the former situation, corresponding to a worst case, the disclosure value is the supremum over the strategies of the probability to disclose the secret (maximisation), whereas in the latter case, the disclosure is the infimum (minimisation). We address quantitative problems (comparing the optimal value with a threshold) and qualitative ones (when the threshold is zero or one) related to both forms of disclosure for a fixed or finite horizon. For all problems, we characterise their decidability status and their complexity. We discover a surprising asymmetry: on the one hand optimal strategies may be chosen among deterministic ones in maximisation problems, while it is not the case for minimisation. On the other hand, for the questions addressed here, more minimisation problems than maximisation ones are decidable.

7.6. D-SPACES: Implementing Declarative Semantics for Spatially Structured Information

Published in [13].

We introduce in this paper D-SPACES, an implementation of constraint systems with space and extrusion operators. Constraint systems are algebraic models that allow for a semantic language-like representation of information in systems where the concept of space is a primary structural feature. We give this information mainly an epistemic interpretation and consider various agents as entities acting upon it. D-SPACES is coded as a c++11 library providing implementations for constraint systems, space functions and extrusion functions. The interfaces to access each implementation are minimal and thoroughly documented. D-SPACES also provides property-checking methods as well as an implementation of a specific type of constraint systems (a boolean algebra). This last implementation serves as an entry point for quick access and proof of concept when using these models. Furthermore, we offer an illustrative example in the form of a small social network where users post their beliefs and utter their opinions.

7.7. Unbounded product-form Petri nets

Published in [14].

Computing steady-state distributions in infinite-state stochastic systems is in general a very difficult task. Product-form Petri nets are those Petri nets for which the steady-state distribution can be described as a natural product corresponding, up to a normalising constant, to an exponentiation of the markings. However, even though some classes of nets are known to have a product-form distribution, computing the normalising constant can be hard. The class of (closed) Π^3-nets has been proposed in an earlier work, for which it is shown that one can compute the steady-state distribution efficiently. However these nets are bounded. In this paper, we generalise queuing Markovian networks and closed Π^3-nets to obtain the class of open Π^3-nets, that generate infinite-state systems. We show interesting properties of these nets: (1) we prove that liveness can be decided in polynomial time, and that reachability in live Π^3-nets can be decided in polynomial time; (2) we show that we can decide ergodicity of such nets in polynomial time as well; (3) we provide a pseudo-polynomial time algorithm to compute the normalising constant.

7.8. Statistical Model-Checking for Autonomous Vehicle Safety Validation

Published in [20].

We present an application of statistical model-checking to the verification of an autonomous vehicle controller. Our goal is to check safety properties in various traffic situations. More specifically, we focus on a traffic jam situation.

The controller is specified by a C++ program. Using sensors, it registers positions and velocities of nearby vehicles and modifies the position and velocity of the controlled vehicle to avoid collisions. We model the environment using a stochastic high level Petri net, where random behaviors of other vehicles can be described. We use HASL, a quantitative variant of linear temporal logic, to express the desired properties. A large family of performance indicators can be specified in HASL and we target in particular the expectation of travelled distance or the collision probability.

We evaluate the properties of this model using COSMOS1. This simulation tool implements numerous statistical techniques such as sequential hypothesis testing and most confidence range computation methods. Its efficiency allowed us to conduct several experiments with success.

7.9. Une sémantique formelle pour les modèles Simulink

Published in [19].

De nombreux projets industriels, notamment dans la construction automobile, font appel à la suite d'outils Simulink pour la conception et la validation de composants critiques représentant des systèmes hybrides c'est-à-dire combinant des aspects discrets et continus. Cependant les formalismes associés ne disposent pas d'une sémantique formelle ce qui peut diminuer la confiance des ingénieurs vis-à-vis des résultats produits. Nous proposons ici une telle sémantique en procédant en deux étapes. Nous développons d'abord une sémantique exacte mais non exécutable. Puis nous l'enrichissons d'une sémantique opérationnelle approchée avec pour objectif une quantification de l'erreur issue de cette approximation.

7.10. The Logical View on Continuous Petri Nets

Published in [5].

Continuous Petri nets are a relaxation of classical discrete Petri nets in which transitions can be fired a fractional number of times, and consequently places may contain a fractional number of tokens. Such continuous Petri nets are an appealing object to study since they over approximate the set of reachable configurations of their discrete counterparts, and their reachability problem is known to be decidable in polynomial time. The starting point of this paper is to show that the reachability relation for continuous Petri nets is definable by a sentence of linear size in the existential theory of the rationals with addition and order. Using this characterization, we obtain decidability and complexity results for a number of classical decision problems for continuous Petri nets. In particular, we settle the open problem about the precise complexity of reachability set inclusion. Finally, we show how continuous Petri nets can be incorporated inside the classical backward coverability algorithm for discrete Petri nets as a pruning heuristic in order to tackle the symbolic state explosion problem. The cornerstone of the approach we present is that our logical characterization enables us to leverage the power of modern SMT-solvers in order to yield a highly performant and robust decision procedure for coverability in Petri nets. We demonstrate the applicability of our approach on a set of standard benchmarks from the literature.

7.11. Memoryless Determinacy of Finite Parity Games: Another Simple Proof

Published in [24].

Memoryless determinacy of (infinite) parity games is an important result with numerous applications. It was first independently established by Emerson and Jutla [1] and Mostowski [2] but their proofs involve elaborate developments. The elegant and simpler proof of Zielonka [3] still requires a nested induction on the finite number of priorities and on ordinals for sets of vertices. There are other proofs for finite games like the one of Bjørklund, Sandberg and Vorobyovin [4] that relies on relating infinite and finite duration games. We present here another simple proof that finite parity games are determined with memoryless strategies using induction on the number of relevant states. The closest proof that relies on induction over non absorbing states is the one of Graedel [5]. However instead of focusing on a single appropriate vertex for induction as we do here, he considers two reduced games per vertex, for all the vertices of the game. The idea of reasoning about a single state has been inspired to me by the analysis of finite stochastic priority games by Karelovic and Zielonka [6].

7.12. Interval iteration algorithm for MDPs and IMDPs

Markov Decision Processes (MDP) are a widely used model including both non-deterministic and probabilistic choices. Minimal and maximal probabilities to reach a target set of states, with respect to a policy resolving non-determinism, may be computed by several methods including value iteration. This algorithm, easy to implement and efficient in terms of space complexity, iteratively computes the probabilities of paths of increasing length. However, it raises three issues: (1) defining a stopping criterion ensuring a bound on the approximation, (2) analysing the rate of convergence, and (3) specifying an additional procedure to obtain the exact values once a sufficient number of iterations has been performed. The first two issues are still open and, for the third one, an upper bound on the number of iterations has been proposed. Based on a graph analysis and transformation of MDPs, we address these problems. First we introduce an interval iteration algorithm, for which the stopping criterion is straightforward. Then we exhibit its convergence rate. Finally we significantly improve the upper bound on the number of iterations required to get the exact values. We extend our approach to also deal with Interval Markov Decision Processes (IMDP) that can be seen as symbolic representations of MDPs.

7.13. Alignment-Based Trace Clustering

Published in [9].

A novel method to cluster event log traces is presented in this paper. In contrast to the approaches in the literature, the clustering approach of this paper assumes an additional input: a process model that describes the current process. The core idea of the algorithm is to use model traces as centroids of the clusters detected, computed from a generalization of the notion of alignment. This way, model explanations of observed behavior are the driving force to compute the clusters, instead of current model agnostic approaches, e.g., which group log traces merely on their vector-space similarity. We believe alignment-based trace clustering provides results more useful for stakeholders. Moreover, in case of log incompleteness, noisy logs or concept drift, they can be more robust for dealing with highly deviating traces. The technique of this paper can be combined with any clustering technique to provide model explanations to the clusters computed. The proposed technique relies on encoding the individual alignment problems into the (pseudo-)Boolean domain, and has been implemented in our tool DarkSider that uses an open-source solver.

7.14. Aligning Modeled and Observed Behavior: A Compromise Between Complexity and Quality

Published in [17].

Certifying that a process model is aligned with the real process executions is perhaps the most desired feature a process model may have: aligned process models are crucial for organizations, since strategic decisions can be made easier on models instead of on plain data. In spite of its importance, the current algorithmic support for computing alignments is limited: either techniques that explicitly explore the model behavior (which may be worst-case exponential with respect to the model size), or heuristic approaches that cannot guarantee a solution, are the only alternatives. In this paper we propose a solution that sits right in the middle in the complexity spectrum of alignment techniques; it can always guarantee a solution, whose quality depends on the exploration depth used and local decisions taken at each step. We use linear algebraic techniques in combination with an iterative search which focuses on progressing towards a solution. The experiments show a clear reduction in the time required for reaching a solution, without sacrificing significantly the quality of the alignment obtained.

7.15. Temporal Reprogramming of Boolean Networks

Published in [15].

Cellular reprogramming, a technique that opens huge opportunities in modern and regenerative medicine, heavily relies on identifying key genes to perturb. Most of computational methods focus on finding mutations to apply to the initial state in order to control which attractor the cell will reach. However, it has been shown, and is proved in this article, that waiting between the perturbations and using the transient dynamics of the system allow new reprogramming strategies. To identify these temporal perturbations, we consider a qualitative model of regulatory networks, and rely on Petri nets to model their dynamics and the putative perturbations.

Our method establishes a complete characterization of temporal perturbations, whether permanent (mutations) or only temporary, to achieve the existential or inevitable reachability of an arbitrary state of the system. We apply a prototype implementation on small models from the literature and show that we are able to derive temporal perturbations to achieve trans-differentiation.

7.16. Goal-Driven Unfolding of Petri Nets

Published in [10].

Unfoldings provide an efficient way to avoid the state-space explosion due to interleavings of concurrent transitions when exploring the runs of a Petri net. The theory of adequate orders allows one to define finite prefixes of unfoldings which contain all the reachable markings. In this paper we are interested in reachability of a single given marking, called the goal. We propose an algorithm for computing a finite prefix of the unfolding of a 1-safe Petri net that preserves all minimal configurations reaching this goal. Our algorithm combines the unfolding technique with on-the-fly model reduction by static analysis aiming at avoiding the exploration of branches which are not needed for reaching the goal. We present some experimental results.

7.17. Euler's Method Applied to the Control of Switched Systems

Hybrid systems are a powerful formalism for modeling and reasoning about cyber-physical systems. They mix the continuous and discrete natures of the evolution of computerized systems. Switched systems are a special kind of hybrid systems, with restricted discrete behaviours: those systems only have finitely many different modes of (continuous) evolution, with isolated switches between modes. Such systems provide a good balance between expressiveness and controllability, and are thus in widespread use in large branches of industry such as power electronics and automotive control. The control law for a switched system defines the way of selecting the modes during the run of the system. Controllability is the problem of (automatically) synthesizing a control law in order to satisfy a desired property, such as safety (maintaining the variables within a given zone) or stabilisation (confinement of the variables in a close neighborhood around an objective point). In order to compute the control of a switched system, we need to compute the solutions of the differential equations governing the modes. Euler's method is the most basic technique for approximating such solutions. We present here an estimation of the Euler's method local error, using the notion of ”one-sided Lispchitz constant” for modes. This yields a general control synthesis approach which can encompass several features such as bounded disturbance and compositionality.

7.18. An Improved Algorithm for the Control Synthesis of Nonlinear Sampled Switched Systems

Published in [6].

A novel algorithm for the control synthesis for nonlinear switched systems is presented in this paper. Based on an existing procedure of state-space bisection and made available for nonlinear systems with the help of guaranteed integration, the algorithm has been improved to be able to consider longer patterns of modes with a better pruning approach. Moreover, the use of guaranteed integration also permits to take bounded perturbations and varying parameters into account. It is particularly interesting for safety critical applications, such as in aeronautical, military or medical fields. The whole approach is entirely guaranteed and the induced controllers are correct-by-design. Some experimentations are performed to show the important gain of the new algorithm.

7.19. Distributed Control Synthesis Using Euler's Method

Published in [22].

In a previous work, we explained how Euler's method for computing approximate solutions of systems of ordinary differential equations can be used to synthesize safety controllers for sampled switched systems. We continue here this line of research by showing how Euler's method can also be used for synthesizing safety controllers in a distributed manner. The global system is seen as an interconnection of two (or more) sub-systems where, for each component, the sub-state corresponding to the other component is seen as an "input"; the method exploits (a variant of) the notions of incremental input-to-state stability (δ-ISS) and ISS Lyapunov function. We illustrate this distributed control synthesis method on a building ventilation example.

7.20. Control Synthesis of Nonlinear Sampled Switched Systems using Euler's Method

Published in [21].

In this paper, we propose a symbolic control synthesis method for nonlinear sampled switched systems whose vector fields are one-sided Lipschitz. The main idea is to use an approximate model obtained from the forward Euler method to build a guaranteed control. The benefit of this method is that the error introduced by symbolic modeling is bounded by choosing suitable time and space discretizations. The method is implemented in the interpreted language Octave. Several examples of the literature are performed and the results are compared with results obtained with a previous method based on the Runge-Kutta integration method.

7.21. Metastability-Aware Memory-Efficient Time-to-Digital Converter

Published in [11].

We propose a novel method for transforming delay- line time-to-digital converters (TDCs) into TDCs that output Gray code without relying on synchronizers. We formally prove that the inevitable metastable memory upsets (Marino, TC'81) do not induce an additional time resolution error. Our modified design provides suitable inputs to the recent metastability-containing sorting networks by Lenzen and Medina (ASYNC'16) and Bund et al. (DATE'17). In contrast, employing existing TDCs would require using thermometer code at the TDC output (followed by conversion to Gray code) or resolving metastability inside the TDC. The former is too restrictive w.r.t. the dynamic range of the TDCs, while the latter loses the advantage of enabling (accordingly much faster) computation without having to first resolve metastability.

Our all-digital designs are also of interest in their own right: they support high sample rates and large measuring ranges at nearly optimal bit-width of the output, yet maintain the original delay-line?s time resolution. No previous approach unifies all these properties in a single device.

7.22. Brief Announcement: Lower Bounds for Asymptotic Consensus in Dynamic Networks

Published in [12].

7.23. Metastability Tolerant Computing

Published in [16].

Synchronization using flip-flop chains imposes a latency of a few clock cycles when transferring data and control signals between clock domains. We propose a design scheme that avoids this latency by performing synchronization as part of state/data computations while guaranteeing that metastability is contained and its effects tolerated (with an acceptable failure probability). We present a theoretical framework for modeling synchronous state machines in the presence of metastability and use it to prove properties that guarantee some form of reliability. Specifically, we show that the inevitable state/data corruption resulting from propagating metastable states can be confined to a subset of computations. Applications that can tolerate certain failures can exploit this property to leverage low-latency and quasi-reliable operation simultaneously. We demonstrate the approach by designing a Network-on-Chip router with zero-latency asynchronous ports and show via simulation that it outperforms a variant with two flip-flop synchronizers at a negligible cost in packet transfer reliability.

8. Partnerships and Cooperations

8.1. National Initiatives

- Thomas Chatain, Stefan Haar , Serge Haddad and Stefan Schwoon are participating in the ANR Project ALGORECELL.
- Matthias Függer participates in the ANR project FREDDA.

-

8.2. International Initiatives

8.2.1. Inria Associate Teams Not Involved in an Inria International Labs

8.2.1.1. LifeForm

Title: Life Sciences need formal Methods !

International Partner (Institution - Laboratory - Researcher):

Newcastle University (United Kingdom) - School of Computing Science - Victor Khomenko

Start year: 2016

See also: http://projects.lsv.ens-cachan.fr/LifeForm/

This project extends an existing cooperation between the MEXICO team and Newcastle University on partial-order based formal methods for concurrent systems. We enlarge the partnership to bioinformatics and synthetic biology. The proposal addresses addresses challenges concerning formal specification, verification, monitoring and control of synthetic biological systems, with use cases conducted in the Center for Synthetic Biology and the Bioeconomy (CSBB) in Newcastle. A main challenge is to create a solid modelling framework based on Petri-net type models that allow for causality analysis and rapid state space exploration for verification, monitoring and control purposes; a potential extension to be investigated concerns the study of attractors and cell reprogramming in Systems Biology.

8.3. International Research Visitors

8.3.1. Visits of International Scientists

Joost-Pieter Katoen, Aachen, spent two weeks with MEXICO.

8.3.1.1. Internships

Aalok Thakkar, 2nd year student from CMI (India), did a two-month research internship on 'Semantics of Mutation Dynamics' under the supervision of Stefan Haar, from May 2nd to July 21st, 2017.

8.3.2. Visits to International Teams

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

- Matthias Függer is general co-chair of *ASYNC 2018*

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

- Laurent Fribourg was a PC member of:
  - 7th International Conference on New Computational Methods for Inverse Problems, Cachan, 2017
  - The 27th International Symposium on Logic-based Program Synthesis and Transformation, Namur, Belgium, 2017
  - Seventh Workshop on Design, Modeling and Evaluation of Cyber Physical Systems, Seoul, 2017.
- Matthias Függer was a member of the PC of DDECS'17.
- Stefan Haar was a member of the PCs of the conferences MSR 2017 and *ACSD 2017* and of the workshop *ATAED 2017*.
- Serge Haddad was a PC member of the *International Workshop on Petri Nets and Software Engineering* (PNSE) 2017 at Zaragoza, Spain, and of the *11th International Conference on Verification and Evaluation of Computer and Communication Systems* (VECOS) 2017 at Montreal, Québec, Canada. He was also a member of the scientific committee of *Ecole d'été temps-réel* (ETR 2017).

9.1.2.2. Reviewer

- Matthias Függer was a reviewer for ICALP, DISC, FMCAD, ICDCN,and OPODIS.
- Stefan Haar was a reviewer for FOSSACS.
- Stefan Schwoon was a reviewer for the following conferences taking place in 2017: STACS, TACAS, ESOP, ATVA, and FSTTCS.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- Stefan Haar is associate editor of the *Journal of Discrete Event Dynamic Systems: Theory and Applications*.

9.1.3.2. Reviewer - Reviewing Activities

- Thomas Chatain was a reviewer for Acta Informatica, Artificial Intelligence and Journal of Discrete Event Dynamic Systems.
- Matthias Függer was a reviewer for MFCS.
- Stefan Haar was a reviewer for MSCS and IEEE Transations on Automatic Control.
- Stefan Schwoon was a reviewer for Fundamenta Informaticae, International Journal on Software Tools for Technology Transfer, and the Journal of Discrete Event Dynamic Systems.

9.1.4. Invited Talks

- Laurent Fribourg gave the following invited talk: "Euler's Method Applied to the Control of Switched Systems", at 15th International Conference on Formal Modelling and Analysis of Timed Systems, Berlin, 2017
- Matthias Függer gave invited talks at the Theory of Hardware seminar in Vienna in February, the Noon seminar at Max-Planck Institute for Informatics in April, and the Distributed Computing Seminar at Labri in November.
- Serge Haddad gave the following invited talks:
  - at Centre Fédéré en Vérification, Bruxelles, Belgique on February 24, 2017, entitled *From Continuous Petri nets to Petri nets and Back*;
  - at LACL, Créteil on February 27, 2017, entitled *Probabilistic Disclosure: Maximisation vs. Minimisation*

  - at the MSR 2017 conference, Marseille, France, on November 16, 2017, entitled *Réseaux de Petri discrets et continus : apports réciproques*.
- Claudine Picaronny gave an invited talk on 'Vérification probabiliste, numérique ou statistique', on the 21th of april 2017 at Alea 17, CIRM Marseille, France

9.1.5. Scientific Expertise

- Serge Haddad was expert for the evaluation of the researcher premiums at University Pierre et Marie Curie

9.1.6. Research Administration

- Laurent Fribourg is a member of
  - Comité Direction of Department Sciences et technologies de l'information et de la communication of Université Paris-Saclay,
  - Bureau of Domaine d'Intérêt Majeur émergent du Réseau Francilien en Sciences Informatiques
- Stefan Haar is the president of Inria's *GTRI-COST* committee for international relations, and the head of the SciLex (Software Reliability and Security) axis of the *LABEX*DIGICOSME, and ipso facto a member of *DIGICOSME*'s executive committee and scientific commission.
- Serge Haddad was the president of the HCERES evaluation committee of the laboratory LIAS, Poitiers.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

- Stefan Haar taught one half of L3 level class on Formal Languages at ENS Paris-Saclay (15 h CM, 22.5 EQTD).
- Serge Haddad is professor at ENS Paris-Saclay. Claudine Picaronny, Thomas Chatain, and Stefan Schwoon are associate professors of the same university.
- Serge Haddad is the head of the Computer Science Department, and Stefan Schwoon is in charge of the L3 formation.
- Claudine Picaronny is a co-director of the ENS Paris-Saclay's Mathematics department and a member of the juries of the 'agrégation interne de Mathématiques' and of the second 'concours de Mathématiques' of ENS Paris-Saclay; she is also the coordinator of the mathematics/computer science examination of E3A, parts MP and MC.
- Matthias Függer is teaching "Initiation à la recherche" at ENS Paris-Saclay.

9.2.2. Supervision

Theses in progress:

Hugues Mandon, ENS Paris-Saclay since October 2016, on *computational models and algorithms for the prediction of Cell Reprogramming Strategies*, co-supervised by Stefan Haar and L. Paulevé (LRI)

Juraj Kolcˇák, , ENS Paris-Saclay since March 2017, on *Unfoldings and Abstract Interpretation for Parametric Biological Regulatory Networks*, co-supervised by Stefan Haar and L. Paulevé (LRI).

Adnane Saoud, Université Paris-Saclay since 2016, jointly supervised by Laurent Fribourg and Antoine Girard (Centrale-Supelec).

Engel Lefaucheux, ENS Paris-Saclay since 2015, *Controlling information in probabilistic systems*, jointly supervised by Nathalie Bertrand (SUMO team) and Serge Haddad

Yann Duplouy, IRT SystemX since 2015, *Application of formal methods to the development of embedded systems for autonomous vehicles*, supervised by Béatrice Bérard and Serge Haddad.

Robert Najvirt (TU Wien, Austrian FWF SIC project), realistic delay models with applica- tions in high-speed and low-power circuits, co-supervised by Matthias Függer and Andreas Steininger.

Martin Perner (TU Wien, Austrian FWF SIC project), clock generation on-chip and formalisms suitable to prove correct VLSI circuits, co-supervised by Matthias Függer and Ulrich Schmid.

Juergen Maier (TU Wien, Austrian FWF SIC project), on realistic delay models with applications in high-speed and low-power circuits, with focus on noise and high-order models, co-supervised by Matthias Függer and with Ulrich Schmid.

9.2.3. Juries

- Laurent Fribourg was a member of the Jury of Irini-Eleftheria Mens's PhD Thesis on “Learning regular languages over large alphabets", defended at University of Grenoble, October 2017.
- Stefan Haar was a reviewer of the thesis by Guillaume Madelaine on 'Simplifications Exactes et Structurelles de Réseaux de Réactions Biologiques', defended on February 28 at Lille University, France.
- Serge Haddad was:
  - reviewer in the jury of Bruno Karelovic on *Quantitative Analysis of Stochastic Systems – Priority Games and Populations of Markov Chains* on July 7 2017, University Paris 7
  - president of the jury of Nicolas David on *Réseaux de Petri à Paramètres Discrets* on October 20 University Nantes
  - reviewer in the jury of Thomas Geffroy on *Vers des outils efficaces pour la vérification de systèmes concurrents* on December 12 2017, University Bordeaux
- Claudine Picaronny was Member of the jury of Pierre Carlier's Thesis on 'Verification of Stochastic Timed Automata', on the 8th of december 2017, Mons, Belgium

10. Bibliography

Major publications by the team in recent years

[1] B. BARBOT, S. HADDAD, C. PICARONNY.*Coupling and Importance Sampling for Statistical Model Checking*, in "Proceedings of the 18th International Conference on Tools and Algorithms for Construction and Analysis of Systems (TACAS'12)", Tallinn, Estonia, C. FLANAGAN, B. KÖNIG (editors), Lecture Notes in Computer Science, Springer, March 2012, vol. 7214, p. 331-346, http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/BHP-tacas12.pdf.

[2] S. HADDAD, J. MAIRESSE, H.-T. NGUYEN.*Synthesis and Analysis of Product-form Petri Nets*, in "Fundamenta Informaticae", 2013, vol. 122, n^{o} 1-2, p. 147-172, https://hal.archives-ouvertes.fr/hal-00925774.

Publications of the year

Articles in International Peer-Reviewed Journal

[3] E. FABRE, L. HÉLOUËT, E. LEFAUCHEUX, H. MARCHAND.*Diagnosability of Repairable Faults*, in "Discrete Event Dynamic Systems", June 2017, https://hal.inria.fr/hal-01646911.

[4] S. HAAR, S. HADDAD, T. MELLITI, S. SCHWOON.*Optimal constructions for active diagnosis*, in "Journal of Computer and System Sciences", February 2017, vol. 83, n^{o} 1, p. 101-120 [*DOI :* 10.1016/J.JCSS.2016.04.007], https://hal.archives-ouvertes.fr/hal-01408047.

[5] S. HADDAD, M. BLONDIN, C. HAASE, A. FINKEL.*The Logical View on Continuous Petri Nets*, in "ACM Transactions on Computational Logic", 2017, vol. 18, n^{o} 3, https://hal.inria.fr/hal-01652793.

[6] A. LE COËNT, J. ALEXANDRE DIT SANDRETTO, A. CHAPOUTOT, L. FRIBOURG.*An Improved Algorithm for the Control Synthesis of Nonlinear Sampled Switched Systems*, in "Formal Methods in System Design", November 2017 [*DOI :* 10.1007/S10703-017-0305-8], https://hal.archives-ouvertes.fr/hal-01399337.

International Conferences with Proceedings

[7] B. BÉRARD, S. HAAR, S. SCHMITZ, S. SCHWOON.*The Complexity of Diagnosability and Opacity Verification for Petri Nets*, in "Petri nets 2017", Zaragoza, Spain, W. VAN DER AALST, E. BEST (editors), Lecture Notes in Computer Science, Springer, 2017, https://hal.inria.fr/hal-01484476.

[8] B. BÉRARD, S. HADDAD, E. LEFAUCHEUX.*Probabilistic Disclosure: Maximisation vs. Minimisation* , in "FSTTCS 2017", Kanpur, India, December 2017 [*DOI :* 10.4230/LIPICS.FSTTXS.2017], https://hal.inria.fr/hal-01618955.

[9] T. CHATAIN, J. CARMONA, B. F. VAN DONGEN.*Alignment-Based Trace Clustering*, in "ER 2017 - 36th International Conference on Conceptual Modeling", Valencia, Spain, ER 2017: Conceptual Modeling, Springer, November 2017, vol. 10650, p. 295-308 [*DOI :* 10.1007/978-3-319-69904-2_24], https://hal.inria.fr/hal-01664235.

[10] T. CHATAIN, L. PAULEVÉ.*Goal-Driven Unfolding of Petri Nets*, in "28th International Conference on Concurrency Theory (CONCUR 2017)", Berlin, Germany, September 2017 [*DOI :* 10.4230/LIPICS.CONCUR.2017.14], https://hal.archives-ouvertes.fr/hal-01392203.

[11] M. FÜGGER, A. KINALI, C. LENZEN, T. POLZER.*Metastability-Aware Memory-Efficient Time-to-Digital Converters*, in "IEEE International Symposium on Asynchronous Circuits and Systems (ASYNC)", San Diego, United States, May 2017, https://hal.archives-ouvertes.fr/hal-01652787.

[12] M. FÜGGER, T. NOWAK, M. SCHWARZ.*Brief Announcement: Lower Bounds for Asymptotic Consensus in Dynamic Networks*, in "31st International Symposium on Distributed Computing (DISC)", Wien, Austria, October 2017, vol. Proceedings of the 31st International Symposium on Distributed Computing (DISC 2017), https://hal.archives-ouvertes.fr/hal-01652851.

[13] S. HAAR, S. PERCHY, F. VALENCIA.*D-SPACES: Implementing Declarative Semantics for Spatially Structured Information*, in "11th International Conference on Semantic Computing", San Diego, California, United States, IEEE ICSC 2017, IEEE, January 2017, vol. 11, https://hal.inria.fr/hal-01328189.

[14] V. JUGÉ, P. BOUYER, S. HADDAD.*Unbounded product-form Petri nets*, in "28th International Conference on Concurrency Theory (CONCUR 2017)", Berlin, Germany, 28th International Conference on Concurrency Theory (CONCUR 2017), September 2017, https://arxiv.org/abs/1708.05847 - 31 pages [*DOI :* 10.4230/LIPICS.CONCUR.2017.31], https://hal.archives-ouvertes.fr/hal-01615257.

[15] H. MANDON, S. HAAR, L. PAULEVÉ.*Temporal Reprogramming of Boolean Networks*, in "CMSB 2017 - 15th conference on Computational Methods for Systems Biology", Darmstadt, Germany, J. FERET, H. KOEPPL (editors), Lecture Notes in Computer Science, Springer International Publishing, September 2017, vol. 10545, p. 179 - 195 [*DOI :* 10.1007/978-3-319-67471-1_11], https://hal.archives-ouvertes.fr/hal-01589251.

[16] G. TARAWNEH, M. FÜGGER, C. LENZEN.*Metastability Tolerant Computing*, in "ASYNC17 - 23rd IEEE International Symposium on Asynchronous Circuits and Systems", San Diego, United States, May 2017, https://hal.archives-ouvertes.fr/hal-01652772.

[17] B. F. VAN DONGEN, J. CARMONA, T. CHATAIN, F. TAYMOURI.*Aligning Modeled and Observed Behavior: A Compromise Between Complexity and Quality*, in "CAiSE 2017 - 29th International Conference on Advanced Information Systems Engineering", Essen, Germany, CAiSE 2017: Advanced Information Systems Engineering, Springer, June 2017, vol. 10253, p. 94-109 [*DOI :* 10.1007/978-3-319-59536-8_7], https://hal.inria.fr/hal-01664240.

National Conferences with Proceeding

[18] N. BERTRAND, S. HADDAD, E. LEFAUCHEUX.*Diagnostic et contrôle de la dégradation des systèmes probabilistes*, in "MSR 2017 - Modélisation des Systèmes Réactifs", Marseille, France, November 2017, https://hal.inria.fr/hal-01618922.

[19] B. BÉRARD, Y. DUPLOUY, S. HADDAD.*Une sémantique formelle pour les modèles Simulink*, in "École d'été Temps Réel 2017", Paris, France, August 2017, https://hal.archives-ouvertes.fr/hal-01578297.

Conferences without Proceedings

[20] B. BARBOT, B. BÉRARD, Y. DUPLOUY, S. HADDAD.*Statistical Model-Checking for Autonomous Vehicle Safety Validation*, in "SIA Simulation Numérique", Montigny-le-Bretonneux, France, Société des Ingénieurs de l'Automobile, March 2017, https://hal.archives-ouvertes.fr/hal-01491064.

[21] A. LE COËNT, F. DE VUYST, L. CHAMOIN, L. FRIBOURG.*Guaranteed control synthesis of nonlinear switched systems using Euler method*, in "3rd International Workshop on Symbolic and Numerical Methods for Reachability Analysis", Uppsala, Sweden, April 2017, https://hal.archives-ouvertes.fr/hal-01583164.

Scientific Books (or Scientific Book chapters)

[22] A. LE COËNT, J. ALEXANDRE DIT SANDRETTO, A. CHAPOUTOT, L. FRIBOURG, F. DE VUYST, L. CHAMOIN.*Distributed control synthesis using Euler's method*, in "International Workshop on Reachability Problems", August 2017 [*DOI :* 10.1007/978-3-319-67089-8_9], https://hal.archives-ouvertes.fr/hal-01593225.

Other Publications

[23] L. FRIBOURG.*Euler's method applied to the control of switched systems*, December 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01670089.

[24] S. HADDAD.*Memoryless Determinacy of Finite Parity Games: Another Simple Proof*, June 2017, working paper or preprint, https://hal.inria.fr/hal-01541508.

[25] A. LE COËNT, L. FRIBOURG, J. VACHER.*Control Synthesis for Stochastic Switched Systems using the Tamed Euler Method*, December 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01670579.

References in notes

[26] S. ABBES, A. BENVENISTE, S. HAAR.*A Petri net model for distributed estimation*, in "Proc. MTNS 2004, Sixteenth International Symposium on Mathematical Theory of Networks and Systems, Louvain (Belgium), ISBN 90-5682-517-8", 2004.

[27] S. AKSHAY, N. BERTRAND, S. HADDAD, L. HELOUET.*The steady-state control problem for Markov decision processes*, in "Qest 2013", Buenos Aires, Argentina, K. R. JOSHI, M. SIEGLE, M. STOELINGA, P. R. D'ARGENIO (editors), Springer, September 2013, vol. 8054, p. 290-304, https://hal.inria.fr/hal-00879355.

[28] R. ALUR, K. ETESSAMI, M. YANNAKAKIS.*Realizability and Verification of MSC Graphs*, in "Theor. Comput. Sci.", 2005, vol. 331, n° 1, p. 97–114.

[29] P. BALDAN, TH. CHATAIN, S. HAAR, B. KÖNIG.*Unfolding-based Diagnosis of Systems with an Evolving Topology*, in "Information and Computation", October 2010, vol. 208, n° 10, p. 1169-1192, http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/BCHK-icomp10.pdf.

[30] P. BALDAN, A. CORRADINI, B. KÖNIG, S. SCHWOON.*McMillan's complete prefix for contextual nets*, in "Transactions on Petri Nets and Other Models of Concurrency", November 2008, vol. 1, p. 199–220, Volume 5100 of Lecture Notes in Computer Science.

[31] P. BALDAN, S. HAAR, B. KOENIG.*Distributed Unfolding of Petri Nets*, in "Proc.FOSSACS 2006", LNCS, Springer, 2006, vol. 3921, p. 126-141, Extended version: Technical Report CS-2006-1. Department of Computer Science, University Ca' Foscari of Venice.

[32] F. BASKETT, K. M. CHANDY, R. R. MUNTZ, F. G. PALACIOS.*Open, Closed, and Mixed Networks of Queues with Different Classes of Customers*, in "J. ACM", April 1975, vol. 22, p. 248–260, http://doi.acm.org/10.1145/321879.321887.

[33] A. BENVENISTE, É. FABRE, S. HAAR.*Markov Nets: Probabilistic Models for distributed and concurrent Systems*, in "IEEE Transactions on Automatic Control", 2003, vol. 48 (11), p. 1936-1950, Extended version: IRISA Research Report 1538.

[34] P. BHATEJA, P. GASTIN, M. MUKUND, K. NARAYAN KUMAR.*Local testing of message sequence charts is difficult*, in "Proceedings of the 16th International Symposium on Fundamentals of Computation Theory (FCT'07)", Budapest, Hungary, E. CSUHAJ-VARJÚ, Z. ÉSIK (editors), Lecture Notes in Computer Science, Springer, August 2007, vol. 4639, p. 76-87 [*DOI :* 10.1007/978-3-540-74240-1_8], http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/BGMN-fct07.pdf.

[35] A. BOUILLARD, S. HAAR, S. ROSARIO.*Critical paths in the Partial Order Unfolding of a Stochastic Petri Net*, in "Proceedings of the 7th International Conference on Formal Modelling and Analysis of Timed Systems (FORMATS'09)", Budapest, Hungary, J. OUAKNINE, F. VAANDRAGER (editors), Lecture Notes in Computer Science, Springer, September 2009, vol. 5813, p. 43-57 [*DOI :* 10.1007/978-3-642-04368-0_6], http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/BHR-formats09.pdf.

[36] P. BOUYER, S. HADDAD, P.-A. REYNIER.*Timed Unfoldings for Networks of Timed Automata*, in "Proceedings of the 4th International Symposium on Automated Technology for Verification and Analysis (ATVA'06)", Beijing, ROC, S. GRAF, W. ZHANG (editors), Lecture Notes in Computer Science, Springer, October 2006, vol. 4218, p. 292-306, http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/BHR-atva06.pdf.

[37] G. CHIOLA, C. DUTHEILLET, G. FRANCESCHINIS, S. HADDAD.*Stochastic Well-Formed Colored Nets and Symmetric Modeling Applications*, in "IEEE Transactions on Computers", November 1993, vol. 42, n^{o} 11, p. 1343-1360, http://www.lsv.ens-cachan.fr/Publis/PAPERS/PS/CDFH-toc93.ps.

[38] R. DEBOUK, D. TENEKETZIS.*Coordinated decentralized protocols for failure diagnosis of discrete-event systems*, in "Journal of Discrete Event Dynamical Systems: Theory and Application", 2000, vol. 10, p. 33–86.

[39] J. ESPARZA, K. HELJANKO.*Unfoldings - A Partial-Order Approach to Model Checking*, EATCS Monographs in Theoretical Computer Science, Springer, 2008.

[40] É. FABRE, A. BENVENISTE, C. JARD, S. HAAR.*Diagnosis of Asynchronous Discrete Event Systems, a Net Unfolding Approach*, in "IEEE Trans. Aut. Control", 2003, vol. 48 (5), p. 714-727.

[41] É. FABRE, A. BENVENISTE, C. JARD, S. HAAR.*Distributed monitoring of concurrent and asynchronous systems*, in "Discrete Event Dynamic Systems: theory and application", 2005, vol. 15 (1), p. 33-84, Preliminary version: Proc. CONCUR 2003, LNCS 2761, pp.1–28, Springer.

[42] S. HAAR, A. BENVENISTE, É. FABRE, C. JARD.*Partial Order Diagnosability Of Discrete Event Systems Using Petri Net Unfoldings*, in "42nd IEEE Conference on Decision and Control (CDC)", 2003.

[43] S. HAAR.*Probabilistic Cluster Unfoldings*, in "Fundamenta Informaticae", 2003, vol. 53 (3-4), p. 281-314.

[44] S. LAFORTUNE, Y. WANG, T.-S. YOO.*Diagnostic Decentralisé Des Systèmes A Evénements Discrets*, in "Journal Europeen des Systèmes Automatisés (RS-JESA)", August 2005, vol. 99, n^{o} 99, p. 95–110.

[45] K. G. LARSEN, P. PETTERSSON, W. YI.*Compositional and symbolic model-checking of real-time systems*, in "Proc. of RTSS 1995", IEEE Computer Society, 1995, p. 76-89.

[46] L. RICKER, K. RUDIE.*Know Means No: Incorporating Knowledge into Discrete-Event Control Systems*, in "IEEE Transactions on Automatic Control", September 2000, vol. 45, n^{o} 9, p. 1656–1668.

[47] L. RICKER, K. RUDIE.*Knowledge Is a Terrible Thing to Waste: Using Inference in Discrete-Event Control Problems*, in "IEEE Transactions on Automatic Control", MarchSeptember 2007, vol. 52, n^{o} 3, p. 428–441.

[48] C. RODRÍGUEZ, S. SCHWOON, V. KHOMENKO.*Contextual Merged Processes*, in "34th International Conference on Applications and Theory of Petri Nets (ICATPN'13)", Italy, Lecture Notes in Computer Science, Springer, 2013, vol. 7927, p. 29-48 [*DOI :* 10.1007/978-3-642-38697-8_3], https://hal.archives-ouvertes.fr/hal-00926202.

[49] H. L. S. YOUNES, R. G. SIMMONS.*Statistical probabilistic model checking with a focus on time-bounded properties*, in "Inf. Comput.", September 2006, vol. 204, p. 1368–1409 [*DOI :* 10.1016/J.IC.2006.05.002], http://dl.acm.org/citation.cfm?id=1182767.1182770.

Project-Team PARIETAL

Modelling brain structure, function and variability based on high-field MRI data.

IN COLLABORATION WITH: CEA Neurospin

IN PARTNERSHIP WITH:
Centre CEA-Saclay

RESEARCH CENTER
Saclay - Île-de-France

THEME
Computational Neuroscience and Medicine

Table of contents

Project-Team PARIETAL

Creation of the Project-Team: 2009 July 01

Keywords:

Computer Science and Digital Science:

A3.3. - Data and knowledge analysis
A3.3.2. - Data mining
A3.3.3. - Big data analysis
A3.4. - Machine learning and statistics
A3.4.1. - Supervised learning
A3.4.2. - Unsupervised learning
A3.4.4. - Optimization and learning
A3.4.5. - Bayesian methods
A3.4.6. - Neural networks
A3.4.7. - Kernel methods
A3.4.8. - Deep learning
A5.3.2. - Sparse modeling and image representation
A5.3.3. - Pattern recognition
A5.9.1. - Sampling, acquisition
A5.9.2. - Estimation, modeling
A5.9.6. - Optimization tools
A6.2.4. - Statistical methods
A6.2.6. - Optimization
A9.2. - Machine learning
A9.3. - Signal analysis

Other Research Topics and Application Domains:

B1.2. - Neuroscience and cognitive science
B1.2.1. - Understanding and simulation of the brain and the nervous system
B1.2.2. - Cognitive science
B2.2.6. - Neurodegenerative diseases
B2.6.1. - Brain imaging

1. Personnel

Research Scientists

Bertrand Thirion [Team leader, Inria, Senior Researcher, HDR]
Philippe Ciuciu [CEA, Researcher, HDR]
Denis Alexander Engemann [Inria, Starting Research Position]
Alexandre Gramfort [Inria, Researcher, from Apr 2017, HDR]
Shan Liu [Inria, Researcher, until Oct 2017]
Gaël Varoquaux [Inria, Researcher]
Demian Wassermann [Inria, Researcher, from Oct 2017]

External Collaborators

Sergul Aydore [Amazon, from Apr 2017]
Joseph Salmon [Institut Telecom ex GET Groupe des Ecoles des Télécommunications , from May 2017]

Technical Staff

Kamalaker Reddy Dadi [CEA, until Sep 2017]
Loïc Esteve [Inria, until Nov 2017]
Ana Luisa Grilo Pinho [Inria]
Olivier Grisel [Inria]
Guillaume Lemaitre [Inria]
Joan Massich Vall [Inria, from Dec 2017]
Joris Van Den Bossche [Inria, from Apr 2017]

PhD Students

Pierre Ablin [Inria, from Apr 2017]
Patricio Cerda Reyes [Inria]
Hamza Cherkaoui [CEA, from Nov 2017]
Jérome-Alexis Chevalier [Inria, from Oct 2017]
Jérome Dockès [Inria]
Elvis Dohmatob [Inria, until Oct 2017]
Loubna El Gueddari [Ministère de l'Enseignement Supérieur et de la Recherche]
Andres Hoyos Idrobo [Inria, until Apr 2017]
Hicham Janati [Inria, from Dec 2017]
Carole Lazarus [CEA]
Mathurin Massias [Inria, from Apr 2017]
Arthur Mensch [Univ Paris-Sud]

Post-Doctoral Fellows

Darya Chyzhyk [Inria]
Joke Durnez [Inria]
Daria La Rocca [CEA]
Antonio Andre Monteiro Manoel [CEA, until Nov 2017]
Mehdi Rahim [Inria, until Oct 2017]

Administrative Assistants

Tiffany Caristan [Inria, until Aug 2017]
Corinne Petitot [Inria, from Oct 2017]

2. Overall Objectives

2.1. Overall Objectives

The Parietal team focuses on mathematical methods for modeling and statistical inference based on neuroimaging data, with a particular interest in machine learning techniques and applications of human functional imaging. This general theme splits into four research axes:

- Modeling for neuroimaging population studies,
- Encoding and decoding models for cognitive imaging,
- Statistical and machine learning methods for large-scale data,
- Compressed-sensing for MRI.

Parietal is also strongly involved in open-source software development in scientific Python (machine learning) and for neuroimaging applications.

3. Research Program

3.1. Inverse problems in Neuroimaging

Many problems in neuroimaging can be framed as forward and inverse problems. For instance, brain population imaging is concerned with the *inverse problem* that consists in predicting individual information (behavior, phenotype) from neuroimaging data, while the corresponding *forward problem* boils down to explaining neuroimaging data with the behavioral variables. Solving these problems entails the definition of two terms: a loss that quantifies the goodness of fit of the solution (does the model explain the data well enough ?), and a regularization scheme that represents a prior on the expected solution of the problem. These priors can be used to enforce some properties on the solutions, such as sparsity, smoothness or being piece-wise constant.

Let us detail the model used in typical inverse problem: Let $\mathbf{X}$ be a neuroimaging dataset as an $(n_{subjects}, n_{voxels})$ matrix, where $n_{subjects}$ and n_{voxels} are the number of subjects under study, and the image size respectively, $\mathbf{Y}$ a set of values that represent characteristics of interest in the observed population, written as $(n_{subjects}, n_{features})$ matrix, where $n_{features}$ is the number of characteristics that are tested, and $\mathbf{w}$ an array of shape $(n_{voxels}, n_{features})$ that represents a set of pattern-specific maps. In the first place, we may consider the columns $\mathbf{Y}_1, .., \mathbf{Y}_{n_{features}}$ of Y independently, yielding $n_{features}$ problems to be solved in parallel:

$$\mathbf{Y}_i = \mathbf{X}\mathbf{w}_i + \epsilon_i, \forall i \in \{1, .., n_{features}\},$$

where the vector contains $\mathbf{w}_i$ is the i^{th} row of $\mathbf{w}$. As the problem is clearly ill-posed, it is naturally handled in a regularized regression framework:

$$\widehat{w}_i = \text{argmin}_{w_i} \|\mathbf{Y}_i - \mathbf{X}\mathbf{w}_i\|^2 + \Psi(\mathbf{w}_i), \tag{10}$$

where Ψ is an adequate penalization used to regularize the solution:

$$\Psi(\mathbf{w}; \lambda_1, \lambda_2, \eta_1, \eta_2) = \lambda_1 \|\mathbf{w}\|_1 + \lambda_2 \|\mathbf{w}\|_2 + \eta_1 \|\nabla \mathbf{w}\|_{2,1} + \eta_2 \|\nabla \mathbf{w}\|_{2,2} \tag{11}$$

with λ_1, λ_2, η_1, $\eta_2 \geq 0$ (this formulation particularly highlights the fact that convex regularizers are norms or quasi-norms). In general, only one or two of these constraints is considered (hence is enforced with a non-zero coefficient):

- When $\lambda_1 > 0$ only (LASSO), and to some extent, when $\lambda_1, \lambda_2 > 0$ only (elastic net), the optimal solution $\mathbf{w}$ is (possibly very) sparse, but may not exhibit a proper image structure; it does not fit well with the intuitive concept of a brain map.
- Total Variation regularization (see Fig. 1) is obtained for ($\eta_1 > 0$ only), and typically yields a piece-wise constant solution. It can be associated with Lasso to enforce both sparsity and sparse variations.
- Smooth lasso is obtained with ($\eta_2 > 0$ and $\lambda_1 > 0$ only), and yields smooth, compactly supported spatial basis functions.

Note that, while the qualitative aspect of the solutions are very different, the predictive power of these models is often very close.

The performance of the predictive model can simply be evaluated as the amount of variance in $\mathbf{Y}_i$ fitted by the model, for each $i \in \{1, .., n_{features}\}$. This can be computed through cross-validation, by *learning*$\widehat{\mathbf{w}}_i$ on some part of the dataset, and then estimating $\|\mathbf{Y}_i - \mathbf{X}\widehat{w}_i\|^2$ using the remainder of the dataset.

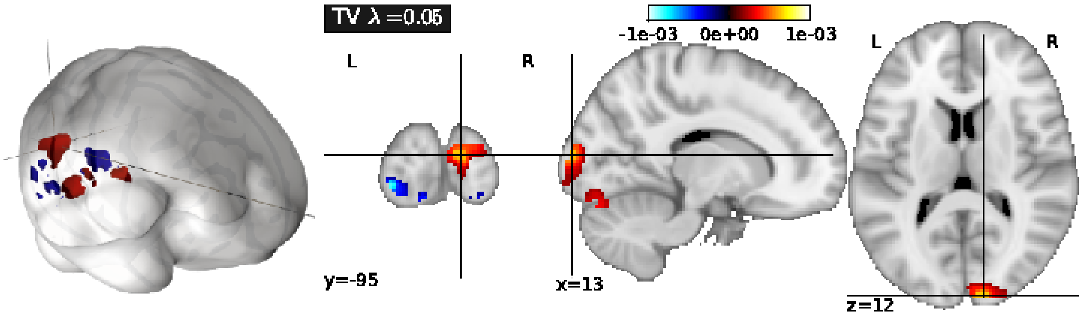


Figure 1. Example of the regularization of a brain map with total variation in an inverse problem. The problem here is to predict the spatial scale of an object presented as a stimulus, given functional neuroimaging data acquired during the presentation of an image. Learning and test are performed across individuals. Unlike other approaches, Total Variation regularization yields a sparse and well-localized solution that also enjoys high predictive accuracy.

This framework is easily extended by considering

- *Grouped penalization*, where the penalization explicitly includes a prior clustering of the features, i.e. voxel-related signals, into given groups. This amounts to enforcing structured priors on the solution.
- *Combined penalizations*, i.e. a mixture of simple and group-wise penalizations, that allow some variability to fit the data in different populations of subjects, while keeping some common constraints.
- *Logistic and hinge regression*, where a non-linearity is applied to the linear model so that it yields a probability of classification in a binary classification problem.
- *Robustness to between-subject variability* to avoid the learned model overly reflecting a few outlying particular observations of the training set. Note that noise and deviating assumptions can be present in both $\mathbf{Y}$ and $\mathbf{X}$
- *Multi-task learning*: if several target variables are thought to be related, it might be useful to constrain the estimated parameter vector $\mathbf{w}$ to have a shared support across all these variables.
  For instance, when one of the variables $\mathbf{Y}_i$ is not well fitted by the model, the estimation of other variables $\mathbf{Y}_j, j \neq i$ may provide constraints on the support of $\mathbf{w}_i$ and thus, improve the prediction of $\mathbf{Y}_i$.

$$\mathbf{Y} = \mathbf{X}\mathbf{w} + \epsilon, \tag{12}$$

then

$$\widehat{w} = \text{argmin}_{\mathbf{w}=(\mathbf{w}_i),i=1..n_f} \sum_{i=1}^{n_f} \|\mathbf{Y_i} - \mathbf{X}\mathbf{w_i}\|^2 + \lambda \sum_{j=1}^{n_{voxels}} \sqrt{\sum_{i=1}^{n_f} \mathbf{w_{i,j}^2}} \tag{13}$$

3.2. Multivariate decompositions

Multivariate decompositions provide a way to model complex data such as brain activation images: for instance, one might be interested in extracting an *atlas of brain regions* from a given dataset, such as regions exhibiting similar activity during a protocol, across multiple protocols, or even in the absence of protocol (during resting-state). These data can often be factorized into spatial-temporal components, and thus can be estimated through *regularized Principal Components Analysis* (PCA) algorithms, which share some common steps with regularized regression.

Let $\mathbf{X}$ be a neuroimaging dataset written as an $(n_{subjects}, n_{voxels})$ matrix, after proper centering; the model reads

$$\mathbf{X} = \mathbf{A}\mathbf{D} + \epsilon, \tag{14}$$

where $\mathbf{D}$ represents a set of n_{comp} spatial maps, hence a matrix of shape (n_{comp}, n_{voxels}), and $\mathbf{A}$ the associated subject-wise loadings. While traditional PCA and independent components analysis are limited to reconstructing components $\mathbf{D}$ within the space spanned by the column of $\mathbf{X}$, it seems desirable to add some constraints on the rows of $\mathbf{D}$, that represent spatial maps, such as sparsity, and/or smoothness, as it makes the interpretation of these maps clearer in the context of neuroimaging. This yields the following estimation problem:

$$\min_{\mathbf{D},\mathbf{A}} \|\mathbf{X} - \mathbf{A}\mathbf{D}\|^2 + \Psi(\mathbf{D}) \text{ s.t. } \|\mathbf{A}_i\| = 1 \; \forall i \in \{1..n_{features}\}, \tag{15}$$

where $(\mathbf{A}_i),\; i \in \{1..n_{features}\}$ represents the columns of $\mathbf{A}$. Ψ can be chosen such as in Eq. (2) in order to enforce smoothness and/or sparsity constraints.
The problem is not jointly convex in all the variables but each penalization given in Eq (2) yields a convex problem on $\mathbf{D}$ for $\mathbf{A}$ fixed, and conversely. This readily suggests an alternate optimization scheme, where $\mathbf{D}$ and $\mathbf{A}$ are estimated in turn, until convergence to a local optimum of the criterion. As in PCA, the extracted components can be ranked according to the amount of fitted variance. Importantly, also, estimated PCA models can be interpreted as a probabilistic model of the data, assuming a high-dimensional Gaussian distribution (probabilistic PCA).

Ultimately, the main limitations to these algorithms is the cost due to the memory requirements: holding datasets with large dimension and large number of samples (as in recent neuroimaging cohorts) leads to inefficient computation. To solve this issue, online methods are particularly attractive [29].

3.3. Covariance estimation

Another important estimation problem stems from the general issue of learning the relationship between sets of variables, in particular their covariance. Covariance learning is essential to model the dependence of these variables when they are used in a multivariate model, for instance to study potential interactions among them and with other variables. Covariance learning is necessary to model latent interactions in high-dimensional observation spaces, e.g. when considering multiple contrasts or functional connectivity data.
The difficulties are two-fold: on the one hand, there is a shortage of data to learn a good covariance model from an individual subject, and on the other hand, subject-to-subject variability poses a serious challenge to the use of multi-subject data. While the covariance structure may vary from population to population, or depending on the input data (activation versus spontaneous activity), assuming some shared structure across problems, such as their sparsity pattern, is important in order to obtain correct estimates from noisy data. Some of the most important models are:

- **Sparse Gaussian graphical models**, as they express meaningful conditional independence relationships between regions, and do improve conditioning/avoid overfit.
- **Decomposable models**, as they enjoy good computational properties and enable intuitive interpretations of the network structure. Whether they can faithfully or not represent brain networks is still an open question.
- **PCA-based regularization of covariance** which is powerful when modes of variation are more important than conditional independence relationships.

Adequate model selection procedures are necessary to achieve the right level of sparsity or regularization in covariance estimation; the natural evaluation metric here is the out-of-sample likelihood of the associated Gaussian model. Another essential remaining issue is to develop an adequate statistical framework to test differences between covariance models in different populations. To do so, we consider different means of parametrizing covariance distributions and how these parametrizations impact the test of statistical differences across individuals.

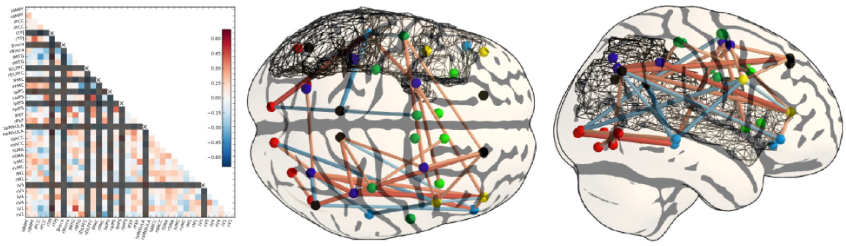


Figure 2. Example of functional connectivity analysis: The correlation matrix describing brain functional connectivity in a post-stroke patient (lesion volume outlined as a mesh) is compared to a group of control subjects. Some edges of the graphical model show a significant difference, but the statistical detection of the difference requires a sophisticated statistical framework for the comparison of graphical models.

4. Application Domains

4.1. Cognitive neuroscience

4.1.1. *Macroscopic Functional cartography with functional Magnetic Resonance Imaging (fMRI)*

The brain as a highly structured organ, with both functional specialization and a complex network organization. While most of the knowledge historically comes from lesion studies and animal electophysiological recordings, the development of non-invasive imaging modalities, such as fMRI, has made it possible to study routinely high-level cognition in humans since the early 90's. This has opened major questions on the interplay between mind and brain , such as: How is the function of cortical territories constrained by anatomy (connectivity) ? How to assess the specificity of brain regions ? How can one characterize reliably inter-subject differences ?

4.1.2. *Analysis of brain Connectivity*

Functional connectivity is defined as the interaction structure that underlies brain function. Since the beginning of fMRI, it has been observed that remote regions sustain high correlation in their spontaneous activity, i.e. in the absence of a driving task. This means that the signals observed during resting-state define a signature of the connectivity of brain regions. The main interest of resting-state fMRI is that it provides easy-to-acquire functional markers that have recently been proved to be very powerful for population studies.

4.1.3. *Modeling of brain processes (MEG)*

While fMRI has been very useful in defining the function of regions at the mm scale, Magneto-encephalography (MEG) provides the other piece of the puzzle, namely temporal dynamics of brain activity, at the ms scale. MEG is also non-invasive. It makes it possible to keep track of precise schedule of mental operations and their interactions. It also opens the way toward a study of the rhythmic activity of the brain. On the other hand, the localization of brain activity with MEG entails the solution of a hard inverse problem.

4.1.4. Current challenges in human neuroimaging (acquisition+analysis)

Human neuroimaging targets two major goals: *i)* the study of neural responses involved in sensory, motor or cognitive functions, in relation to models from cognitive psychology, i.e. the identification of neurophysiological and neuroanatomical correlates of cognition; *ii)* the identification of markers in brain structure and function of neurological or psychiatric diseases. Both goals have to deal with a tension between

- the search for higher spatial [0] resolution to increase **spatial specificity** of brain signals, and clarify the nature (function and structure) of brain regions. This motivates efforts for high-field imaging and more efficient acquisitions, such as compressed sensing schemes, as well as better source localization methods from M/EEG data.
- the importance of inferring brain features with **population-level** validity, hence, contaminated with high variability within observed cohorts, which blurs the information at the population level and ultimately limits the spatial resolution of these observations.

Importantly, the signal-to-noise ratio (SNR) of the data remains limited due to both resolution improvements [0] and between-subject variability. Altogether, these factors have led to realize that results of neuroimaging studies were **statistically weak**, i.e. plagued with low power and leading to unreliable inference [54], and particularly so due to the typically number of subjects included in brain imaging studies (20 to 30, this number tends to increase [55]): this is at the core of the *neuroimaging reproducibility crisis*. This crisis is deeply related to a second issue, namely that only few neuroimaging datasets are publicly available, making it impossible to re-assess a posteriori the information conveyed by the data. Fortunately, the situation improves, lead by projects such as NeuroVault or OpenfMRI. A framework for integrating such datasets is however still missing.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. ERC

Demian Wassermann obtained an ERC starting grant, Neurolang, *Accelerating Neuroscience Research by Unifying Knowledge Representation and Analysis Through a Domain Specific Language*.

Besides, Alexandre Gramfort joined Parietal just after the start of his ERC grant entitled SLAB, *Signal processing and Learning Applied to Brain data*.

6. New Software and Platforms

6.1. Mayavi

FUNCTIONAL DESCRIPTION: Mayavi is the most used scientific 3D visualization Python software. Mayavi can be used as a visualization tool, through interactive command line or as a library. It is distributed under Linux through Ubuntu, Debian, Fedora and Mandriva, as well as in PythonXY and EPD Python scientific distributions. Mayavi is used by several software platforms, such as PDE solvers (fipy, sfepy), molecule visualization tools and brain connectivity analysis tools (connectomeViewer).

- Contact: Gaël Varoquaux
- URL: http://mayavi.sourceforge.net/

[0] and to some extent, temporal, but for the sake of simplicity we focus here on spatial aspects.

[0] The SNR of the acquired signal is proportional to the voxel size, hence an improvement by a factor of 2 in image resolution along each dimension is payed by a factor of 8 in terms of SNR.

6.2. MedInria

KEYWORDS: Visualization - DWI - Health - Segmentation - Medical imaging
SCIENTIFIC DESCRIPTION: It aims at creating an easily extensible platform for the distribution of research algorithms developed at Inria for medical image processing. This project has been funded by the D2T (ADT MedInria-NT) in 2010, renewed in 2012. A fast-track ADT was awarded in 2017 to transition the software core to more recent dependencies and study the possibility of a consortium creation.The Visages team leads this Inria national project and participates in the development of the common core architecture and features of the software as well as in the development of specific plugins for the team's algorithm.
FUNCTIONAL DESCRIPTION: MedInria is a free software platform dedicated to medical data visualization and processing.

- Participants: Maxime Sermesant, Olivier Commowick and Théodore Papadopoulo
- Partners: HARVARD Medical School - IHU - LIRYC - NIH
- Contact: Olivier Commowick
- URL: http://med.inria.fr

6.3. Nilearn

NeuroImaging with scikit learn
KEYWORDS: Health - Neuroimaging - Medical imaging
FUNCTIONAL DESCRIPTION: NiLearn is the neuroimaging library that adapts the concepts and tools of scikit-learn to neuroimaging problems. As a pure Python library, it depends on scikit-learn and nibabel, the main Python library for neuroimaging I/O. It is an open-source project, available under BSD license. The two key components of NiLearn are i) the analysis of functional connectivity (spatial decompositions and covariance learning) and ii) the most common tools for multivariate pattern analysis. A great deal of efforts has been put on the efficiency of the procedures both in terms of memory cost and computation time.

- Participants: Alexandre Abraham, Alexandre Gramfort, Bertrand Thirion, Elvis Dohmatob, Fabian Pedregosa Izquierdo, Gaël Varoquaux, Loïc Estève, Michael Eickenberg and Virgile Fritsch
- Contact: Bertrand Thirion
- URL: http://nilearn.github.io/

6.4. PyHRF

KEYWORDS: Health - Brain - IRM - Neurosciences - Statistic analysis - FMRI - Medical imaging
FUNCTIONAL DESCRIPTION: As part of fMRI data analysis, PyHRF provides a set of tools for addressing the two main issues involved in intra-subject fMRI data analysis : (i) the localization of cerebral regions that elicit evoked activity and (ii) the estimation of the activation dynamics also referenced to as the recovery of the Hemodynamic Response Function (HRF). To tackle these two problems, PyHRF implements the Joint Detection-Estimation framework (JDE) which recovers parcel-level HRFs and embeds an adaptive spatio-temporal regularization scheme of activation maps.

- Participants: Aina Frau Pascual, Christine Bakhous, Florence Forbes, Jaime Eduardo Arias Almeida, Laurent Risser, Lotfi Chaari, Philippe Ciuciu, Solveig Badillo, Thomas Perret and Thomas Vincent
- Partners: CEA - NeuroSpin
- Contact: Florence Forbes
- URL: http://pyhrf.org

6.5. Scikit-learn

KEYWORDS: Regession - Clustering - Learning - Classification - Medical imaging

SCIENTIFIC DESCRIPTION: Scikit-learn is a Python module integrating classic machine learning algorithms in the tightly-knit scientific Python world. It aims to provide simple and efficient solutions to learning problems, accessible to everybody and reusable in various contexts: machine-learning as a versatile tool for science and engineering.
FUNCTIONAL DESCRIPTION: Scikit-learn can be used as a middleware for prediction tasks. For example, many web startups adapt Scikitlearn to predict buying behavior of users, provide product recommendations, detect trends or abusive behavior (fraud, spam). Scikit-learn is used to extract the structure of complex data (text, images) and classify such data with techniques relevant to the state of the art.

Easy to use, efficient and accessible to non datascience experts, Scikit-learn is an increasingly popular machine learning library in Python. In a data exploration step, the user can enter a few lines on an interactive (but non-graphical) interface and immediately sees the results of his request. Scikitlearn is a prediction engine . Scikit-learn is developed in open source, and available under the BSD license.

- Participants: Alexandre Gramfort, Bertrand Thirion, Fabian Pedregosa Izquierdo, Gaël Varoquaux, Loïc Estève, Michael Eickenberg and Olivier Grisel
- Partners: CEA - Logilab - Nuxeo - Saint Gobain - Tinyclues - Telecom Paris
- Contact: Olivier Grisel
- URL: http://scikit-learn.org

6.6. MODL

Massive Online Dictionary Learning
KEYWORDS: Pattern discovery - Machine learning
FUNCTIONAL DESCRIPTION: Matrix factorization library, usable on very large datasets, with optional sparse and positive factors.

- Participants: Arthur Mensch, Gaël Varoquaux, Bertrand Thirion and Julien Mairal
- Contact: Arthur Mensch
- Publications: Subsampled online matrix factorization with convergence guarantees - Stochastic Subsampling for Factorizing Huge Matrices
- URL: http://github.com/arthurmensch/modl

6.7. MNE

MNE-Python
KEYWORDS: Neurosciences - EEG - MEG - Signal processing - Machine learning
FUNCTIONAL DESCRIPTION: Open-source Python software for exploring, visualizing, and analyzing human neurophysiological data: MEG, EEG, sEEG, ECoG, and more.

- Contact: Alexandre Gramfort
- URL: http://martinos.org/mne/

7. New Results

7.1. Joint prediction of multiple scores captures better individual traits from brain images

To probe individual variations in brain organization, population imaging relates features of brain images to rich descriptions of the subjects such as genetic information or behavioral and clinical assessments. Capturing common trends across these measurements is important: they jointly characterize the disease status of patient groups. In particular, mapping imaging features to behavioral scores with predictive models opens the way toward more precise diagnosis. Here we propose to jointly predict all the dimensions (behavioral scores) that make up the individual profiles, using so-called multi-output models. This approach often boosts prediction accuracy by capturing latent shared information across scores. We demonstrate the efficiency of multi-output models on two independent resting-state fMRI datasets targeting different brain disorders (Alzheimer's Disease and schizophrenia). Furthermore, the model with joint prediction generalizes much better to a new cohort: a model learned on one study is more accurately transferred to an independent one. Finally, we show how multi-output models can easily be extended to multi-modal settings, combining heterogeneous data sources for a better overall accuracy.

More information can be found in Fig. 3 in [46].

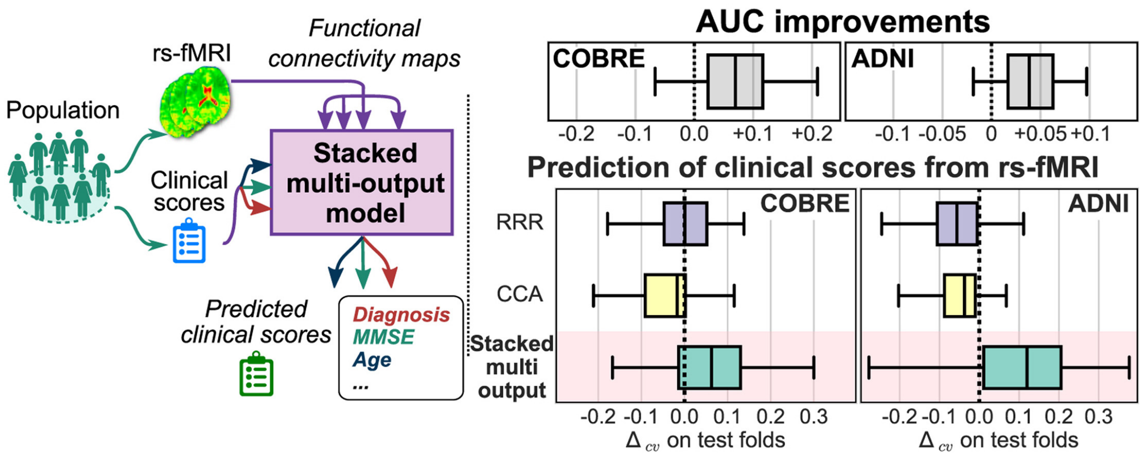


Figure 3. Joint prediction of multiple scores captures better individual traits from brain images

7.2. Population-shrinkage of covariance to estimate better brain functional connectivity

Brain functional connectivity, obtained from functional Magnetic Resonance Imaging at rest (r-fMRI), reflects inter-subject variations in behavior and characterizes neuropathologies. It is captured by the covariance matrix between time series of remote brain regions. With noisy and short time series as in r-fMRI, covariance estimation calls for penalization, and shrinkage approaches are popular. Here we introduce a new covariance estimator based on a non-isotropic shrinkage that integrates prior knowledge of the covariance distribution over a large population. The estimator performs shrinkage tailored to the Riemannian geometry of symmetric positive definite matrices, coupled with a probabilistic modeling of the subject and population covariance distributions. Experiments on a large-scale dataset show that such estimators resolve better intra-and inter-subject functional connectivities compared existing co-variance estimates. We also demonstrate that the estimator improves the relationship across subjects between their functional-connectivity measures and their behavioral assessments. More information can be found in Fig. 4 in [47].

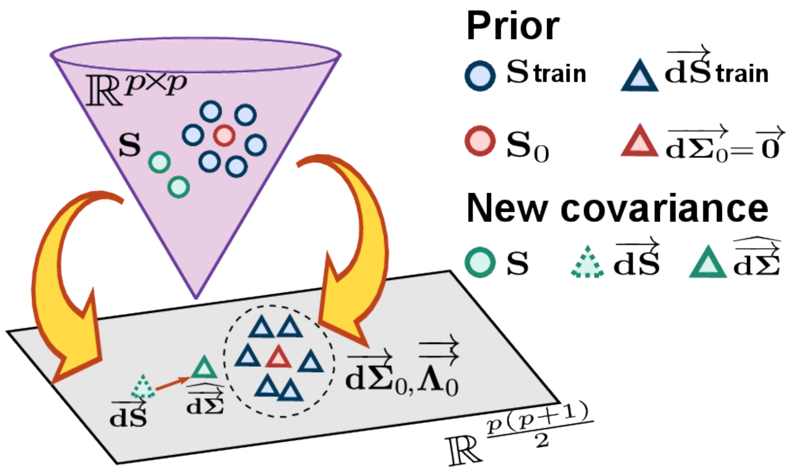


Figure 4. (a) Shrunk embedding estimation workflow: the empirical covariance is estimated from r-fMRI time-series; it is projected onto a tangent space built from a prior population; the embedding is then shrunk towards the prior $(\overrightarrow{d\Sigma_0}, \overrightarrow{\Lambda}_0)$. *(b) Principle of tangent embedding shrinkage towards population distribution.*

7.3. Fast Regularized Ensembles of Models

Brain decoding relates behavior to brain activity through predictive models. These are also used to identify brain regions involved in the cognitive operations related to the observed behavior. Training such multivariate models is a high-dimensional statistical problem that calls for suitable priors. State of the art priors –eg small total-variation– enforce spatial structure on the maps to stabilize them and improve prediction. However, they come with a hefty computational cost. We build upon very fast dimension reduction with spatial structure and model ensembling to achieve decoders that are fast on large datasets and increase the stability of the predictions and the maps. Our approach, fast regularized ensemble of models (FReM), includes an implicit spatial regularization by using a voxel grouping with a fast clustering algorithm. In addition, it aggregates different estimators obtained across splits of a cross-validation loop, each time keeping the best possible model. Experiments on a large number of brain imaging datasets show that our combination of voxel clustering and model ensembling improves decoding maps stability and reduces the variance of prediction accuracy. Importantly, our method requires less samples than state-of-the-art methods to achieve a given level of prediction accuracy. Finally, FreM is highly parallelizable, and has lower computation cost than other spatially-regularized methods.

More information can be found in Fig. 5 in [23].

7.4. time decoding

Most current functional Magnetic Resonance Imaging (fMRI) decoding analyses rely on statistical summaries of the data resulting from a deconvolution approach: each stimulation event is associated with a brain response. This standard approach leads to simple learning procedures, yet it is ill-suited for decoding events with short

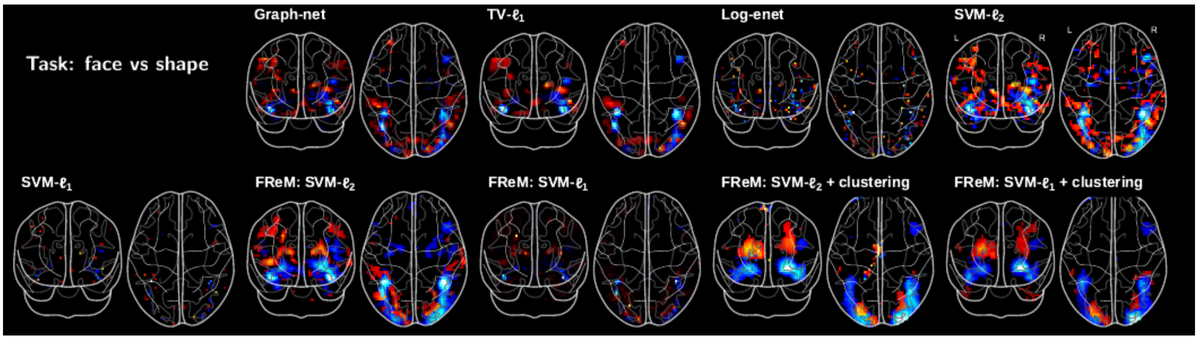


Figure 5. Qualitative comparison of decoder weight maps: Weight maps for different discriminative tasks on the HCP dataset. The maps are thresholded at the 99 percentile for visualization purposes. These correspond to a face-recognition task. The weight maps obtained with TV-L1 and FReM methods with clustering outline the organization of the functional areas of the visual mosaic, such as: primary visual areas, lateral occipital complex, the face and place specific regions in the fusiform gyrus.

inter-stimulus intervals. In order to overcome this issue, we propose a novel framework that separates the spatial and temporal components of the prediction by decoding the fMRI time-series continuously, i.e. scan-by-scan. The stimulation events can then be identified through a deconvolution of the reconstructed time series. We show that this model performs as well as or better than standard approaches across several datasets, most notably in regimes with small inter-stimuli intervals (3 to 5s), while also offering predictions that are highly interpretable in the time domain. This opens the way toward analyzing datasets not normally thought of as suitable for decoding and makes it possible to run decoding on studies with reduced scan time.

More information can be found in Fig. 6 in [28].

7.5. Hierarchical Region-Network Sparsity for High-Dimensional Inference in Brain Imaging

Structured sparsity penalization has recently improved statistical models applied to high-dimensional data in various domains. As an extension to medical imaging, the present work incorporates priors on network hierarchies of brain regions into logistic-regression to distinguish neural activity effects. These priors bridge two separately studied levels of brain architecture: functional segregation into regions and functional integration by networks. Hierarchical region-network priors are shown to better classify and recover 18 psychological tasks than other sparse estimators. Varying the relative importance of region and network structure within the hierarchical tree penalty captured complementary aspects of the neural activity patterns. Local and global priors of neurobiological knowledge are thus demonstrated to offer advantages in generalization performance, sample complexity, and domain interpretability.

More information can be found in Fig. 7 in [48].

7.6. Cross-validation failure: small sample sizes lead to large error bars

Predictive models ground many state-of-the-art developments in statistical brain image analysis: decoding, MVPA, searchlight, or extraction of biomarkers. The principled approach to establish their validity and usefulness is cross-validation, testing prediction on unseen data. Here, we raise awareness on error bars of cross-validation, which are often underestimated. Simple experiments show that sample sizes of many neuroimaging studies inherently lead to large error bars, eg $\pm 10\%$ for 100 samples. The standard error

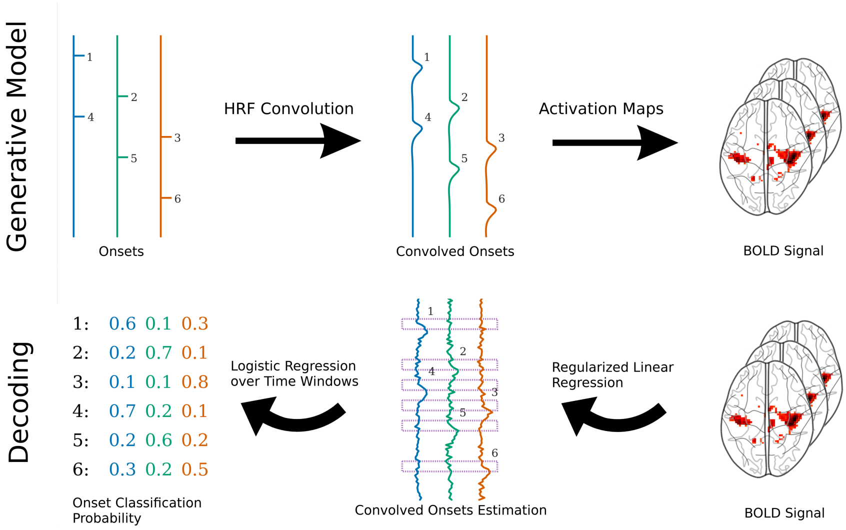


Figure 6. Schema of the time-domain decoding model. Straight arrows represent generative steps, while curved ones represent estimation steps.

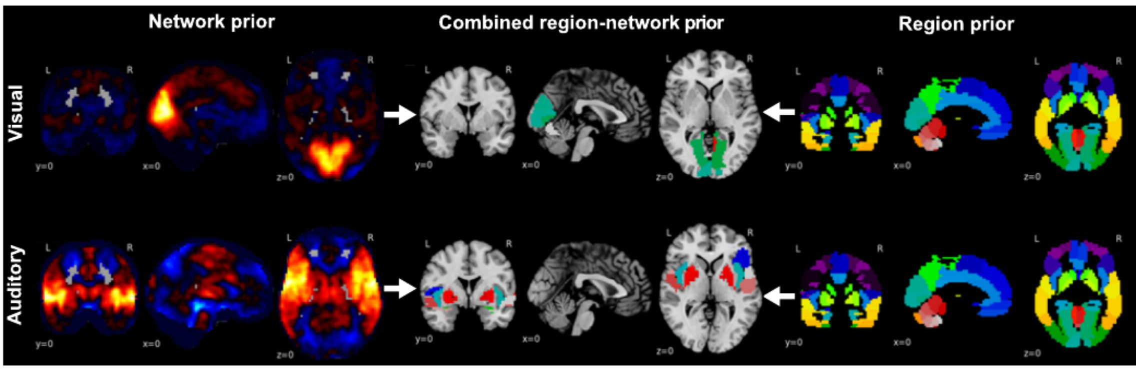


Figure 7. Building blocks of the hierarchical region-network tree. Displays the a-priori neurobiological knowledge introduced into the classification problem by hierarchical structured sparsity. Left: Continuous, partially overlapping brain network priors (hot-colored) accommodate the functional integration perspective of brain organization. Right: Discrete, non-overlapping brain region priors (single-colored) accommodate the functional segregation perspective. Middle: These two types of predefined voxel groups are incorporated into a joint hierarchical prior of parent networks with their descending region child nodes. Top to bottom: Two exemplary region-network priors are shown, including the early cortices that process visual and sound information from the environment.

across folds strongly underestimates them. These large error bars compromise the reliability of conclusions drawn with predictive models, such as biomarkers or methods developments where, unlike with cognitive neuroimaging MVPA approaches, more samples cannot be acquired by repeating the experiment across many subjects. Solutions to increase sample size must be investigated, tackling possible increases in heterogeneity of the data.

More information can be found in Fig. 8 in [33].

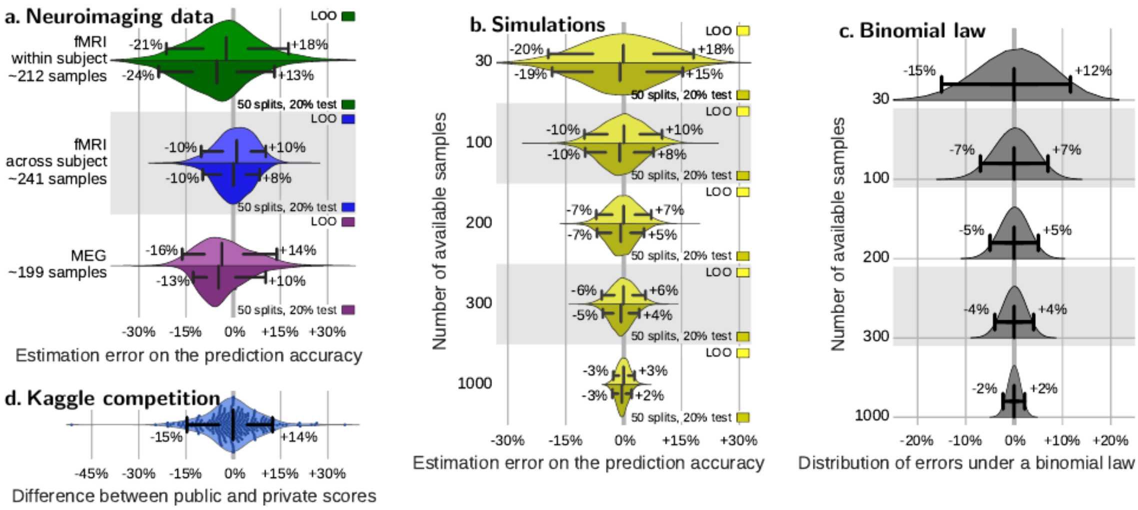


Figure 8. Cross-validation errors. a – Distribution of errors between the prediction accuracy as assessed via cross-validation (average across folds) and as measured on a large independent test set for different types of neuroimaging data. b – Distribution of errors between the prediction accuracy as assessed via cross-validation on data of various sample sizes and as measured on 10 000 new data points for simple simulations. c – Distribution of errors as given by a binomial law: difference between the observed prediction error and the population value of the error, p = 75%, for different sample sizes. d – Discrepancies between private and public score. Each dot represents the difference between the accuracy of a method on the public test data and the private one. The scores are retrieved from http://www.kaggle.com/c/mlsp-2014-mri, in which 144 subjects were used total, 86 for training predictive model, 30 for the public test set, and 28 for the private test set. The bar and whiskers indicate the median and the 5th and 95th percentile. Measures on cross-validation (a and b) are reported for two reasonable choices of cross-validation strategy: leave one out (leave one run out or leave one subject out in data with multiple runs or subjects), or 50-times repeated splitting of 20% of the data.

7.7. Autoreject: Automated artifact rejection for MEG and EEG data

We present an automated algorithm for unified rejection and repair of bad trials in magnetoencephalography (MEG) and electroencephalography (EEG) signals. Our method capitalizes on cross-validation in conjunction with a robust evaluation metric to estimate the optimal peak-to-peak threshold – a quantity commonly used for identifying bad trials in M/EEG. This approach is then extended to a more sophisticated algorithm which estimates this threshold for each sensor yielding trial-wise bad sensors. Depending on the number of bad sensors, the trial is then repaired by interpolation or by excluding it from subsequent analysis. All steps of the algorithm are fully automated thus lending itself to the name Autoreject. In order to assess the practical significance of the algorithm, we conducted extensive validation and comparisons with state-of-the-art methods on four public datasets containing MEG and EEG recordings from more than 200 subjects. The comparisons include purely qualitative efforts as well as quantitatively benchmarking against human supervised and semi-automated preprocessing pipelines. The algorithm allowed us to automate the preprocessing of MEG data from the Human Connectome Project (HCP) going up to the computation of the evoked responses. The automated nature of our method minimizes the burden of human inspection, hence supporting scalability and reliability demanded by data analysis in modern neuroscience.

More information can be found in Fig. 9 and in [24].

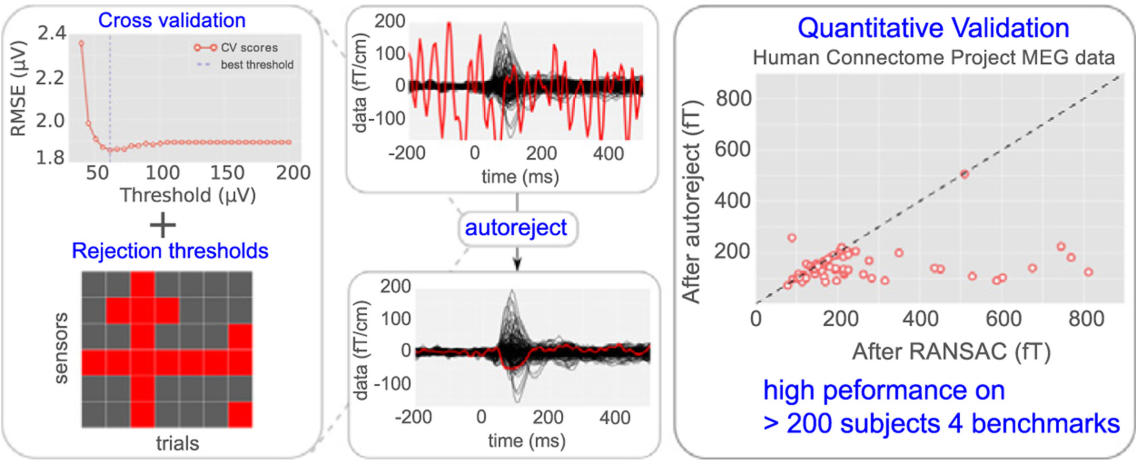


Figure 9. Autoreject: Automated artifact rejection for MEG and EEG data.

7.8. Learning Neural Representations of Human Cognition across Many fMRI Studies

Cognitive neuroscience is enjoying rapid increase in extensive public brain-imaging datasets. It opens the door to large-scale statistical models. Finding a unified perspective for all available data calls for scalable and automated solutions to an old challenge: how to aggregate heterogeneous information on brain function into a universal cognitive system that relates mental operations/cognitive processes/psychological tasks to brain networks? We cast this challenge in a machine-learning approach to predict conditions from statistical brain maps across different studies. For this, we leverage multi-task learning and multi-scale dimension reduction to learn low-dimensional representations of brain images that carry cognitive information and can be robustly associated with psychological stimuli. Our multi-dataset classification model achieves the best prediction performance on several large reference datasets, compared to models without cognitive-aware low-dimension representations; it brings a substantial performance boost to the analysis of small datasets, and can be introspected to identify universal template cognitive concepts.

More information can be found in Fig. 10 in [45].

7.9. SPARKLING: Novel Non-Cartesian Sampling Schemes for Accelerated 2D Anatomical Imaging at 7T Using Compressed Sensing

We have presented for the first time the implementation of non-Cartesian trajectories on a 7T scanner for 2D anatomical imaging. The proposed SPARKLING curves (Segmented Projection Algorithm for Random K-space sampLING) are a new type of non-Cartesian segmented sampling trajectories which allow fast and

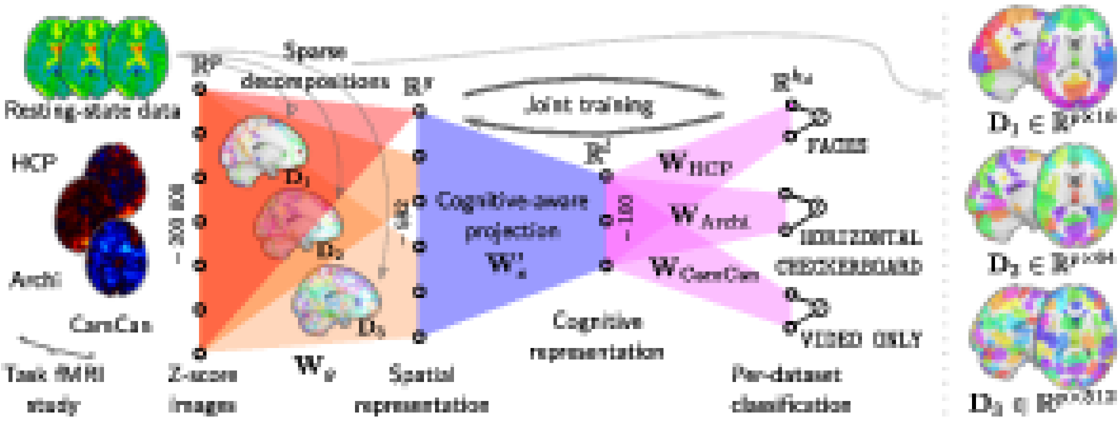


Figure 10. ***Model architecture: Three-layer multi-dataset classification.*** *The first layer (orange) is learned from data acquired outside of cognitive experiments and captures a spatially coherent signal at multiple scales, the second layer (blue) embeds these representations in a space common to all datasets, from which the conditions are predicted (pink) from multinomial models.*

efficient coverage of the k-space according to a chosen variable density. To demonstrate their potential, a high-resolution (0.4×0.4×3.0mm 3) T2*-weighted image was acquired with an 8-fold undersampled SPARKLING trajectory. Images were reconstructed using non-linear iterative reconstructions derived from the Compressed Sensing theory.

More information can be found in Fig. 11 in [43], [42].

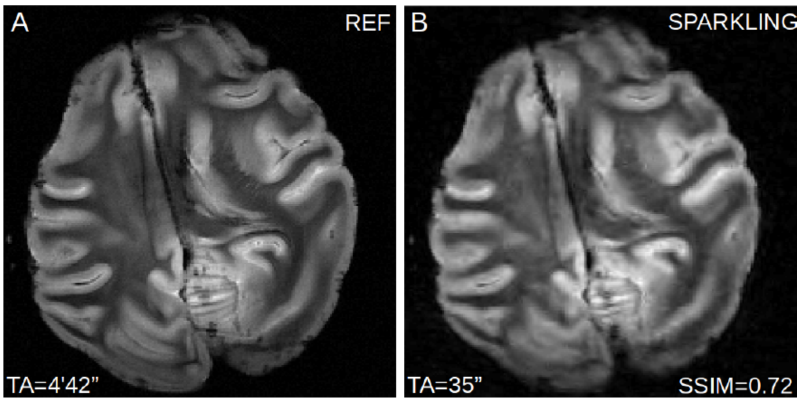


Figure 11. T2-weighted Gradient-Echo transversal slice of the ex vivo baboon brain is displayed for A) the full Cartesian reference sampling lasting 4 minutes and 42 seconds and B) the presented 8-fold undersampled SPARKLING trajectories lasting only 35 seconds. A good visual quality of the subsampled reconstruction was observed with the preservation of major folded patterns. The structural similarity score between the SPARKLING reconstructions and the reference was also very satisfactory.*

7.10. A Projection Method on Measures Sets

We consider the problem of projecting a probability measure π on a set $\mathcal{M}_N$ of Radon measures. The projection is defined as a solution of the following variational problem:

$$\inf_{\mu \in \mathcal{M}_N} \|h * (\mu - \pi)\|_2^2,$$

where $h \in L^2(\Omega)$ is a kernel, $\Omega \subset \mathbb{R}^d$ and denotes the convolution operator. To motivate and illustrate our study, we show that this problem arises naturally in various practical image rendering problems such as stippling (representing an image with N dots) or continuous line drawing (representing an image with a continuous line). We provide a necessary and sufficient condition on the sequence $(\mathcal{M}_N)_{N \in \mathbb{N}}$ that ensures weak convergence of the projections $(\mu_N^*)_{N \in \mathbb{N}}$ to π. We then provide a numerical algorithm to solve a discretized version of the problem and show several illustrations related to computer-assisted synthesis of artistic paintings/drawings.

More information can be found in [19].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Grants with Industry

8.1.1. The Wendelin FUI project

The Wendelin project has been granted on December 3rd, 2014. It has been selected at the *Programme d'Investissements d'Avenir (PIA)* that supports "cloud computing et Big Data". It gives visibility and fosters the French technological big data sector, and in particular the scikit-learn library, the NoSQL “NEO” et the decentralized “SlapOS” cloud, three open-source software supported by the Systematic *pôle de compétitivité*.

Scikit-learn is a worldwide reference library for machine learning. Gaël Varoquaux, Olivier Grisel and Alexandre Gramfort have been major players in the design of the library and Scikit-learn has then been supported by the growing scientific Python community. It is currently used by major internet companies as well as dynamic start-ups, including Google, Airbnb, Spotify, Evernote, AWeber, TinyClues; it wins more than half of the data science "Kaggle" competitions. Scikit-learn makes it possible to predict future outcomes given a training data, and thus to optimize company decisions. Almost 1 million euros will be invested to improve the algorithmic core of scikit-learn through the Wendelin project thanks to the Inria, ENS and Institut Mines Télécom teams. In particular, scikit-learn will be extended in order to ease online prediction and to include recent stochastic gradient algorithms.

NEO is the native NoSQL base of the Python language. It was initially designed by Nexedi and is currently used and embedded in the main software of company information systems. More than one million euros will be invested into NEO, so that scikit-learn can process within 10 years (out-of-core) data of 1 exabyte size.

Paris13 university and the Mines Télécom institute will extend the SlapOS distributed mesh cloud to deploy Wendelin in *Big Data as a Service* (BDaaS) mode, to achieve the interoperability between the Grid5000 and Teralab infrastructures and to extend the cloud toward smart sensor systems.

The combination of scikit-learn, NEO and SlapOS will improve the predictive maintenance of industrial plants with two major use cases: connected windmills (GDF SUEZ, Woelfel) and customer satisfaction in car sale systems (MMC Rus). In both cases it is about non-personal, yet profitable big data. The Wendelin project actually demonstrates that Big data can improve infrastructure and everyday-life equipment without intrusive data collection. For more information, please see http://www.wendelin.io.

The project partners are:

- Nexedi (leader)
- GDF SUEZ
- Abilian
- 2ndQuadrant
- Institut Mines Télécom
- Inria
- Université Paris 13

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. CoSmic project

Participants: Philippe Ciuciu [Correspondant], Carole Lazarus, Loubna El Gueddari.

This is a collaborative project with Jean-Luc Stark, (CEA) funded by the DRF-impulsion CEA program.

Compressed Sensing is a recent theory in maths that allows the perfect recovery of signals or images from compressive acquisition scenarios. This approach has been popularized in MRI over the last decade as well as in astrophysics (noticeably in radio-astronomy). So far, both of these fields have developed skills in CS separately. The aim of the COSMIC project is to foster collaborations between CEA experts in MRI (Parietal team within NeuroSpin) and in astrophysics (CosmoStat lab within the Astrophysics Department). These interactions will allow us to share different expertise in order to improve image quality, either in MRI or in radio-astronomy (thanks to the interferometry principle). In this field, given the data delivered by radio-telescopes, the goal consists in extracting high temporal resolution information in order to study fast transient events.

9.1.2. BrainAMP project

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux, Antonio Andre Monteiro Manoel.

This is a collaborative project with Lenka Zdeborová, Theoretical Physics Institute (CEA) funded by the DRF-impulsion CEA program.

In many scientific fields, the data acquisition devices have benefited of hardware improvement to increase the resolution of the observed phenomena, leading to ever larger datasets. While the dimensionality has increased, the number of samples available is often limited, due to physical or financial limits. This is a problem when these data are processed with estimators that have a large sample complexity, such as multivariate statistical models. In that case it is very useful to rely on structured priors, so that the results reflect the state of knowledge on the phenomena of interest. The study of the human brain activity through high-field MRI belongs among these problems, with up to 10^6 features, yet a set of observations limited by cost and participant comfort.

We are missing fast estimators for multivariate models with structured priors, that furthermore provide statistical control on the solution. Approximate message passing methods are designed to work optimally with low-sample-complexity, they accommodate rather generic class of priors and come with an estimation of statistical significance. They are therefore well suited for our purposes.

We want to join forces to design a new generation of inverse problem solvers that can take into account the complex structure of brain images and provide guarantees in the low-sample-complexity regime. To this end, we will first adapt AMP to the brain mapping setting, using first standard sparsity priors (e.g. Gauss-Bernoulli) on the model. We will then consider more complex structured priors that control the variation of the learned image patterns in space. Crucial gains are expected from the use of the EM algorithm for parameter setting, that comes naturally with AMP. We will also examine the estimators provided by AMP for statistical significance. BrainAMP will design a reference inference toolbox released as a generic open source library. We expect a 3- to 10-fold improvement in CPU time, that will benefit to large-scale brain mapping investigations.

9.1.3. iConnectom project

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux, Elvis Dohmatob.

This is a Digiteo project (2014-2017).

Mapping brain functional connectivity from functional Magnetic Resonance Imaging (MRI) data has become a very active field of research. However, analysis tools are limited and many important tasks, such as the empirical definition of brain networks, remain difficult due to the lack of a good framework for the statistical modeling of these networks. We propose to develop population models of anatomical and functional connectivity data to improve the alignment of subjects brain structures of interest while inferring an average template of these structures. Based on this essential contribution, we will design new statistical inference procedures to compare the functional connections between conditions or populations and improve the sensitivity of connectivity analysis performed on noisy data. Finally, we will test and validate the methods on multiple datasets and distribute them to the brain imaging community.

9.1.4. MetaCog project

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux, Jérome Dockès.

This is a Digicosme project (2016-2019) and a collaboration with Fabian Suchanek (Telecom Paritech).

Understanding how cognition emerges from the billions of neurons that constitute the human brain is a major open problem in science that could bridge natural science –biology– to humanities –psychology. Psychology studies performed on humans with functional Magnetic Resonance Imaging (fMRI) can be used to probe the full repertoire of high-level cognitive functions. While analyzing the resulting image data for a given experiment is a relatively well-mastered process, the challenges in comparing data across multiple datasets poses serious limitation to the field. Indeed, such comparisons require to pool together brain images acquired under different settings and assess the effect of different *experimental conditions* that correspond to psychological effects studied by neuroscientists.

Such meta-analyses are now becoming possible thanks to the development of public data resources –OpenfMRI http://openfmri.org and NeuroVault http://neurovault.org. As many others, researchers of the Parietal team understand these data sources well and contribute to them. However, in such open-ended context, the description of experiments in terms of cognitive concepts is very difficult: there is no universal definition of cognitive terms that could be employed consistently by neuroscientists. Hence meta-analytic studies loose power and specificity. On the other hand, http://brainspell.org provide a set of curated annotation, albeit on much less data, that can serve as a seed or a ground truth to define a consensual ontology of cognitive concepts. Relating these terms to brain activity poses another challenge, of statistical nature, as brain patterns form high-dimensional data in perspective with the scarcity and the noise of the data.

The purpose of this project is to learn a semantic structure in cognitive terms from their occurrence in brain activation. This structure will simplify massive multi-label statistical-learning problems that arise in brain mapping by providing compact representations of cognitive concepts while capturing the imprecision on the definition these concepts.

9.1.5. HighDimStat project

Participants: Bertrand Thirion [Correspondant], Jérome-Alexis Chevalier, Joseph Salmon.

This is a Digicosme project (2017-2020) and a collaboration with Joseph Salmon (Telecom Paritech).

The HiDimStat project aims at handling uncertainty in the challenging context of high dimensional regression problem. Though sparse models have been popularized in the last twenty years in contexts where many features can explain a phenomenon, it remains a burning issue to attribute confidence to the predictive models that they produce. Such a question is hard both from the statistical modeling point of view, and from a computation perspective. Indeed, in practical settings, the amount of features at stake (possibly up to several millions in high resolution brain imaging) limit the application of current methods and require new algorithms to achieve computational efficiency. We plan to leverage recent developments in sparse convex solvers as well as more efficient reformulations of testing and confidence interval estimates to provide several communities with practical software handling uncertainty quantification. Specific validation experiments will be performed in the field of brain imaging.

9.1.6. AMPHI project

Participants: Bertrand Thirion [Correspondant], Joseph Salmon, Antonio Andre Monteiro Manoel.

This is a Digicosme project (2017-2020) and a collaboration with Joseph Salmon (Telecom Paritech) and Lenka Zdeborova (CEA, IPhT).

In many scientific fields, the data acquisition devices have benefited of hardware improvement to increase the resolution of the observed phenomena, leading to ever larger datasets. While the dimensionality has increased, the number of samples available is often limited, due to physical or financial limits. This is a problem when these data are processed with estimators that have a large sample complexity, such as multivariate statistical models. In that case it is very useful to rely on structured priors, so that the results reflect the state of knowledge on the phenomena of interest. The study of the human brain activity through neuroimaging belongs among these problems, with up to 10^6 features, yet a set of observations limited by cost and participant comfort. We are missing fast estimators for multivariate models with structured priors, that furthermore provide statistical control on the solution. Approximate message passing (AMP) methods are designed to work optimally with low- sample-complexity, they accommodate rather generic class of priors and come with an estimation of statistical significance. They are therefore well suited for our purposes. We want to join forces to design a new generation of inverse problem solvers that can take into account the complex structure of brain images and provide guarantees in the low-sample-complexity regime. To this end, we will first adapt AMP to the brain mapping setting, using first standard sparsity priors (e.g. Gauss-Bernoulli) on the model. We will then consider more complex structured priors that control the variation of the learned image patterns in space. Crucial gains are expected from the use of the EM algorithm for parameter setting, that comes naturally with AMP. We will also examine the estimators provided by AMP for statistical significance. AMPHI will design a reference inference toolbox released as a generic open source library. We expect a 3- to 10-fold improvement in CPU time, that will benefit to large-scale brain mapping investigations.

9.1.7. CDS2

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux, Guillaume Lemaitre, Joris Van Den Bossche.

CDS2 is an "Strategic research initiative” of the Paris Saclay University Idex http://datascience-paris-saclay.fr. Although it groups together many partners of the Paris Saclay ecosystem, Parietal has been deeply involved in the project. It currently funds a post-doc for Guillaume Lemaitre and an engineer position for Joris van den Bossche.

9.2. National Initiatives

9.2.1. ANR

9.2.1.1. MultiFracs project

Participants: Philippe Ciuciu [Correspondant], Daria La Rocca.

The scale-free concept formalizes the intuition that, in many systems, the analysis of temporal dynamics cannot be grounded on specific and characteristic time scales. The scale-free paradigm has permitted the relevant analysis of numerous applications, very different in nature, ranging from natural phenomena (hydrodynamic turbulence, geophysics, body rhythms, brain activity,...) to human activities (Internet traffic, population, finance, art,...).

Yet, most successes of scale-free analysis were obtained in contexts where data are univariate, homogeneous along time (a single stationary time series), and well-characterized by simple-shape local singularities. For such situations, scale-free dynamics translate into global or local power laws, which significantly eases practical analyses. Numerous recent real-world applications (macroscopic spontaneous brain dynamics, the central application in this project, being one paradigm example), however, naturally entail large multivariate data (many signals), whose properties vary along time (non-stationarity) and across components (non-homogeneity), with potentially complex temporal dynamics, thus intricate local singular behaviors.

These three issues call into question the intuitive and founding identification of scale-free to power laws, and thus make uneasy multivariate scale-free and multifractal analyses, precluding the use of univariate methodologies. This explains why the concept of scale-free dynamics is barely used and with limited successes in such settings and highlights the overriding need for a systematic methodological study of multivariate scale-free and multifractal dynamics. The Core Theme of MULTIFRACS consists in laying the theoretical foundations of a practical robust statistical signal processing framework for multivariate non homogeneous scale-free and multifractal analyses, suited to varied types of rich singularities, as well as in performing accurate analyses of scale-free dynamics in spontaneous and task-related macroscopic brain activity, to assess their natures, functional roles and relevance, and their relations to behavioral performance in a timing estimation task using multimodal functional imaging techniques.

This overarching objective is organized into 4 Challenges:

1. Multivariate scale-free and multifractal analysis,
2. Second generation of local singularity indices,
3. Scale-free dynamics, non-stationarity and non-homogeneity,
4. Multivariate scale-free temporal dynamics analysis in macroscopic brain activity.

9.2.1.2. Niconnect project

Participants: Bertrand Thirion, Gaël Varoquaux [Correspondant], Kamalaker Reddy Dadi, Darya Chyzhyk, Mehdi Rahim.

- **Context:** The NiConnect project (2012-2017) arises from an increasing need of medical imaging tools to diagnose efficiently brain pathologies, such as neuro-degenerative and psychiatric diseases or lesions related to stroke. Brain imaging provides a non-invasive and widespread probe of various features of brain organization, that are then used to make an accurate diagnosis, assess brain rehabilitation, or make a prognostic on the chance of recovery of a patient. Among different measures extracted from brain imaging, functional connectivity is particularly attractive, as it readily probes the integrity of brain networks, considered as providing the most complete view on brain functional organization.
- **Challenges:** To turn methods research into popular tool widely usable by non specialists, the NiConnect project puts specific emphasis on producing high-quality open-source software. NiConnect addresses the many data analysis tasks that extract relevant information from resting-state fMRI datasets. Specifically, the scientific difficulties are *i)* conducting proper validation of the models and tools, and *ii)* providing statistically controlled information to neuroscientists or medical doctors. More importantly, these procedures should be robust enough to perform analysis on limited quality data, as acquiring data on diseased populations is challenging and artifacts can hardly be controlled in clinical settings.
- **Outcome of the project:** In the scope of computer science and statistics, NiConnect pushes forward algorithms and statistical models for brain functional connectivity. In particular, we are investigating structured and multi-task graphical models to learn high-dimensional multi-subject brain connectivity models, as well as spatially-informed sparse decompositions for segmenting structures from brain imaging. With regards to neuroimaging methods development, NiConnect provides systematic comparisons and evaluations of connectivity biomarkers and a software library embedding best-performing state-of-the-art approaches. Finally, with regards to medical applications, the NiConnect project also plays a support role in on going medical studies and clinical trials on neurodegenerative diseases.
- **Consortium**
  - Parietal Inria research team: applied mathematics and computer science to model the brain from MRI
  - LIF INSERM research team: medical image data analysis and modeling for clinical applications

– CATI center: medical image processing center for large scale brain imaging studies
– Henri-Mondor hospital neurosurgery and neuroradiology: clinical teams conducting research on treatments for neurodegenerative diseases, in particular Huntington and Parkinson diseases
– Logilab: consulting in scientific computing

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

9.3.1.1. HBP

Title: The Human Brain Project

Programm: FP7

Duration: October 2013 - September 2016

Coordinator: EPFL

Partners:

- Inria contact: Olivier Faugeras
- Understanding the human brain is one of the greatest challenges facing 21st century science. If we can rise to the challenge, we can gain profound insights into what makes us human, develop new treatments for brain diseases and build revolutionary new computing technologies. Today, for the first time, modern ICT has brought these goals within sight. The goal of the Human Brain Project, part of the FET Flagship Programme, is to translate this vision into reality, using ICT as a catalyst for a global collaborative effort to understand the human brain and its diseases and ultimately to emulate its computational capabilities. The Human Brain Project will last ten years and will consist of a ramp-up phase (from month 1 to month 36) and subsequent operational phases. This Grant Agreement covers the ramp-up phase. During this phase the strategic goals of the project will be to design, develop and deploy the first versions of six ICT platforms dedicated to Neuroinformatics, Brain Simulation, High Performance Computing, Medical Informatics, Neuromorphic Computing and Neurorobotics, and create a user community of research groups from within and outside the HBP, set up a European Institute for Theoretical Neuroscience, complete a set of pilot projects providing a first demonstration of the scientific value of the platforms and the Institute, develop the scientific and technological capabilities required by future versions of the platforms, implement a policy of Responsible Innovation, and a programme of transdisciplinary education, and develop a framework for collaboration that links the partners under strong scientific leadership and professional project management, providing a coherent European approach and ensuring effective alignment of regional, national and European research and programmes. The project work plan is organized in the form of thirteen subprojects, each dedicated to a specific area of activity. A significant part of the budget will be used for competitive calls to complement the collective skills of the Consortium with additional expertise.

9.4. International Initiatives

9.4.1. Inria Associate Teams Not Involved in an Inria International Labs

9.4.1.1. MetaMRI

Title: Machine learning for meta-analysis of functional neuroimaging data

International Partner (Institution - Laboratory - Researcher):

Stanford (United States) - Department of Psychology - Russ Poldrack

Start year: 2015

See also: https://team.inria.fr/metamri

Neuroimaging produces huge amounts of complex data that are used to better understand the relations between brain structure and function. Observing that the neuroimaging community is still largely missing appropriate tools to store and organize the knowledge related to the data, Parietal team and Poldrack's lab, have decided to join forces to set up a framework for functional brain image meta-analysis, i.e. a framework in which several datasets can be jointly analyzed in order to accumulate information on the functional specialization of brain regions. MetaMRI will build upon Poldrack's lab expertise in handling, sharing and analyzing multi-protocol data and Parietal's recent developments of machine learning libraries to develop a new generation of meta-analytic tools.

9.4.1.2. LargeBrainNets

Title: Characterizing Large-scale Brain Networks Using Novel Computational Methods for dMRI and fMRI-based Connectivity

International Partner (Institution - Laboratory - Researcher):

Stanford (United States) - Stanford Cognitive and Systems Neuroscience Laboratory - Vinod Menon

Start year: 2016

See also: http://www-sop.inria.fr/members/Demian.Wassermann/large-brain-nets.html

In the past two decades, brain imaging of neurotypical individuals and clinical populations has primarily focused on localization of function and structures in the brain, revealing activation in specific brain regions during performance of cognitive tasks through modalities such as functional MRI. In parallel, technologies to identify white matter structures have been developed using diffusion MRI. More recently, interest has shifted towards developing a deeper understanding of the brain's intrinsic architecture and its influence on cognitive and affective information processing. Using for this resting state fMRI and diffusion MRI to build the functional and structural networks of the human brain.

The human brain is a complex patchwork of interconnected regions, and graph-theoretical approaches have become increasingly useful for understanding how functionally connected systems engender, and constrain, cognitive functions. The functional nodes of the human brain and their structural inter-connectivity, collectively the "connectome", are, however, poorly understood. Critically, there is a dearth of computational methods for reliably identifying functional nodes of the brain and their structural inter-connectivity in vivo, despite an abundance of high-quality data from the Human Connectome Project (HCP). Devising and validating methods for investigating the human connectome has therefore taken added significance.

The first major goal of this project is to develop and validate appropriate sophisticated computational and mathematical tools for identifying functional nodes at the whole-brain level and measuring structural and functional connectivity between them, using state-of-the-art human brain imaging techniques and open-source HCP data. To this end, we will first develop and validate novel computational tools for (1) identifying stable functional nodes of the human brain using resting-state functional MRI and (2) measuring structural connectivity between functional nodes of the brain using multi-shell high-angular diffusion MRI. Due to the complementarity of the two imaging techniques fMRI and dMRI, our novel computational methods methods, the synergy between the two laboratories of this associate team will allow us to reveal in unprecedented detail the structural and functional connectivity of the human brain.

The second major goal of this project is to use our newly developed computational tools to characterize normal structural and functional brain networks in neurotypical adults.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

Parietal has welcome François Meyer, Univ Colorado at Boulder, for a six months visit (Jan-June 2017), funded by a D'Alembert fellowship of Paris Saclay University. The project of François is to assess novel statistical models of functional connectivity based on the generalized resistivity model he has developed within a graph theoretical framework.

9.5.2. Visits to International Teams

9.5.2.1. Research Stays Abroad

9.5.2.1.1. Denis Engemann

has spent two months in Boston (April-May) with the MEG Core lab, Athinoula A. Martinos Center (MGH/Harvard-MIT) working on functional connectivity methods and population analysis for MEG.

9.5.2.1.2. Arthur Mensch

has spent 3 months in Japan (Sept-Dec) with NTT, working on dynamic time warping problems with Mathieu Blondel.

9.5.2.1.3. Jérome Dockès

has spent two months with Poldracklab at Stanford, as part of the MetaMRI associated team. He has worked on the statistical relationships between neuroscientific concepts (whether anatomical or cognitive) and brain activation loci.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Selection

10.1.1.1. Member of the Conference Program Committees

10.1.1.1.1. Gaël Varoquaux

MICCAI, NIPS, IPMI, ICML, CCN

10.1.1.1.2. Bertrand Thirion

NIPS, IPMI, OHBM, PRNI

10.1.2. Journal

10.1.2.1. Member of the Editorial Boards

- Gaël Varoquaux, Editor, Neuroimage
- Alexandre Gramfort, Editor, Frontiers in Brain Imaging Methods, Journal of Machine Learning Research (JMLR)
- Bertrand Thirion, Editor, Frontiers in Brain Imaging Methods

10.1.2.2. Reviewer - Reviewing Activities

10.1.2.2.1. Gaël Varoquaux

Nature Methods, JMLR, PLOS Comp Bio, NeuroImage, IEEE TBME, Annals of Applied Statistics, Biological Psychiatry, MedIA

10.1.2.2.2. Alexandre Gramfort

JMLR, PLOS Comp Bio, NeuroImage, IEEE TBME, IEEE TMI, IEEE TSP, MedIA, Brain Topography, NIPS, ICML, ICLR.

10.1.2.2.3. Philippe Ciuciu

PLOS One, NeuroImage, HBM, IEEE TMI, IEEE TSP, IEEE SP letters, MedIA, J of Neuroscience, JCBFM, ISBI, ICASSP, EUSIPCO.

10.1.2.2.4. Denis Engemann

PLOS Biology, PLOS Computational Biology, Nature Scientific Reports, Neuroimage, Neuroimage Clinical, Human Brain Mapping, Journal of Machine Learning Research, Brain Topography, Brain Connectivity, Journal of Alzheimer's Disease, PLOS ONE, Frontiers in Neuroscience, Psychiatry and Clinical Neurosciences, Sensors

10.1.2.2.5. Bertrand Thirion

Nature communications, Neuroimage, Medical Image Analysis, IEEE TMI, PNAS, PLOS Comp Bio.

10.1.3. Invited Talks

- Gaël Varoquaux, Keynote, BDEC Workshop 2017, Wuxi, China
- Gaël Varoquaux, Keynote, scipy 2017, Austin TX
- Gaël Varoquaux, Keynote, Swiss Python summit 2017, Zurich, Switzerland
- Gaël Varoquaux, Keynote, NIPS learning with limited labels workshop, Long Beach, CA
- Philippe Ciuciu, Seminar in the neuroscience Department, NYU, School of Medicine, May 2017, NYC.
- Philippe Ciuciu, Laufer Center Seminar Series, May 2017, Stony Brook Univ, NY.
- Philippe Ciuciu, Dedale workshop on Dictionary learning on manifolds, Sep. 2017, Nice, France
- Philippe Ciuciu, Colloquium at Perform centre, Dec 2017, Montreal, Canada
- Philippe Ciuciu, Seminar at Ecole Polytechnique Montreal, Dec 2017, Montreal, Canada
- Alexandre Gramfort, 2017 EU-US Frontiers of Engineering Symposium, UC Davis, CA
- Alexandre Gramfort, Computational Challenges in Image Processing, Turing Institute, Cambridge Univ., UK
- Alexandre Gramfort, Machine learning for functional brain imaging Symposium and Workshop, Karolinska Institute, Sweden
- Denis A. Engemann, Lessons learned from high-dimensional statistics and M/EEG: From automated preprocessing to detection of consciousness, University of Naples, Italy, March 2017
- Denis A. Engemann, Learning from Oscillations – New Vistas for Translational Neuroimaging, MIT, Cambridge MA, April 2017
- Denis A. Engemann, Building better biomarkers using M/EEG and statistical learning, Eespo University, Finland, June 2017
- Bertrand Thirion, ARSEP, February 2017, Paris
- Bertrand Thirion, AI for Medical Imaging symposium, June 2017, Paris
- Bertrand Thirion, Causality and big data workshop, October 2017, Saclay
- Bertrand Thirion, Machine learning workshop, Telcom Paris Sud, October 2017, Evry
- Bertrand Thirion, Icube seminary, Strasburg, July 2017
- Bertrand Thirion, Laboratoire de Mathématiques d'Orsay, December 2017

10.1.4. Scientific Expertise

- Philippe Ciuciu, Member of the "ANR evualuation committee (CES 45) in Mathématique, informatique, automatique, traitement du signal"
- Philippe Ciuciu, Member of the IEEE BISP committe for the international symposium in biomedical imaging

- Philippe Ciuciu, Member of the Biomedical Imaging & Signal Analytics special area team (SAT) in Eurasip
- Bertrand Thirion, Member of the prospective group of ITMO Neurosciences

10.1.5. Research Administration

- Gaël Varoquaux, Member of the Inria Saclay CDT, "Comission de développement téchnologique"
- Gaël Varoquaux, Member of the Inria Saclay CSD, "Comission de suivi doctoral"
- Gaël Varoquaux, Member of the Inria Saclay cluster committee
- Alexandre Gramfort, Member of the "comité de pilotage" DataIA
- Alexandre Gramfort, Member of the "comité de pilotage" DIM RFSI
- Philippe Ciuciu, Member of the "Inria Saclay scientific committee"
- Philippe Ciuciu, Member of the CEA/DRF Impulsion evaluation committee
- Philippe Ciuciu, Member of the "jury du prix de thèse de la société EEA et du GDR ISIS"
- Bertrand Thirion, Deputy scientific director of Inria Saclay research center
- Bertrand Thirion, Leader of the Datasense research axis of the Digicosme Labex
- Bertrand Thirion, Member of the steering committee of the Dataia Convergence Institute
- Bertrand Thirion, Member of the steering committee of the Computer Science Deprtment of paris Saclay University.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

10.2.1.1. Gaël Varoquaux

Master : Machine learning with scikit-learn, 2 heures équivalent TD, M1, ENSAE, France

Master : Brain functional connectivity, 8 heures équivalent TD, M2, Télécom Paris Tech, France

Doctorat : Machine learning for brain imaging with nilearn, 16 heures équivalent TD, Université de Montréal, Canada

Doctorat : Ecole d'hiver, Computational Brain Connectivity Mapping, Juan-Les-Pins, 1h30

Doctorat : Ecole d'hiver, Computational Brain Connectivity Mapping, 1h30, Juan-Les-Pins

Doctorat : Machine learning for brain imaging, 1h00, IPAM, UCLA, Los Angeles, USA

Doctorat : Scipy: numerical algorithms in Python, 1h30, Euroscipy 2017, Erlangen, Germany

Doctorat : Estimation of brain functional connectomes, 30mn, OHBM 2017, Vancouver, Canada

Doctorat : Software engineering for reproducible science, Open science summer school 2017, EPFL, Lausanne, Switzerland

10.2.1.2. Philippe Ciuciu

Master 2 : "Functional MRI: From data acquisition to analysis", 3h,Univ. Paris V René Descartes & Télécom-Paristech, Master of Biomedical Engineering

Master 2 : "FMRI data analysis", 3h, Univ. Paris-Saclay, Master of medical Physics

10.2.1.3. Alexandre Gramfort

Master 2 : "Functional Brain Imaging with EEG and MEG", 6h, Univ. Paris V René Descartes & Télécom-Paristech, Master of Biomedical Engineering

Master 2 : "Optimization for Data Science", 20h, Univ. Paris-Saclay, Master of Mathematics / Data Science

Master 2 : "Data Camp", 18h, Univ. Paris-Saclay, Master of Mathematics / Data Science

Ecole d'hiver, Computational Brain Connectivity Mapping, Juan-Les-Pins, 1h30

Scikit-learn tutorial, Scipy Conf., Austin, USA, 6h

MNE tutorial, Univ. Libre de Bruxelles, Belgique, 6h

10.2.1.4. Denis A. Engemann

MNE-Python workshop, University of Bari, June 2017, 14h

10.2.1.5. Bertrand Thirion

Master : Functional Neuroimaging and brain-computer interfaces, 12h, MVA, ENS Paris-Saclay, France

Doctorat : OHBM course 2017, Pattern Recognition for brain Imaging, Vancouver, Canada

Master: Resting-State Functional Magnetic Resonance Imaging, 3h, Paris V, France.

10.2.2. Supervision

PhD in progress : Patricio Cerda, Statistical methods for analysis across datasets, started 01/09/2016, supervisors: Balazs Kegl (CNRS, LAL), Gaël Varoquaux

PhD in progress : Jérome Dockès, Mining text and brain activity to learning the semantics of cognitive science, started 01/09/2016, supervisors: Fabian Suchanek (Télécom Paristech), Gaël Varoquaux

PhD in progress : Arthur Mensch, Learning representations of fMRI, started 01/09/2015, supervisors: Bertrand Thirion, Gaël Varoquaux, Julien Mairal (Inria Grenoble)

PhD: Elvis Dohmatob, Amélioration de connectivité fonctionnelle par utilisation de modèles déformables dans l'estimation de décompositions spatiales des images de cerveau, defended sept 2017, supervisors Bertrand Thirion, Gaël Varoquaux

10.2.2.1. Philippe Ciuciu

PhD in progress: Carole Lazarus (3rd year, Director)

PhD in progress: Loubna El Gueddari (2nd year, Director)

PhD in progress: Hamza Cherkaoui (1st year, Director)

PhD in progress: Sylvain Lannuzel (1st year, co-director)

10.2.2.2. Alexandre Gramfort

PhD defended: Albert Thomas

PhD defended: Romain Laby

PhD in progress: Yousra Bekhti

PhD in progress: Mainak Jas

PhD in progress: Tom Dupré La Tour

PhD in progress: Stanislas Chambon

PhD in progress: Pierre Ablin

PhD in progress: Mathurin Massias

PhD in progress: Hicham Janati

10.2.2.3. Denis A. Engemann

PhD in progress: Mainak Jas (co-supervision on 1 main project)

PhD in progress: Sami aboud (co-supervision on 1 main project)

PhD in progress: Kamalakar Reddy (co-supervision)

10.2.2.4. Bertrand Thirion

PhD defended: Elvis Dohmatob (co-supervision with G.Varoquaux)

PhD in progress: Arthur Mensch (co-supervision with G. Varoquaux and J. Mairal)

PhD in progress: Jérome Dockès (co-supervision with G. Varoquaux and F. Suchanek)

PhD in progress: Kamalakar Reddy (co-supervision with G.Varoquaux and D. Engemann)

PhD in progress: Jérome-Alexis Chevalier (co-supervision with J.Salmon)

10.2.3. Juries

10.2.3.1. Philippe Ciuciu

12/15: Reviewer of Lucie Thiebaut-Lonjaret PhD Thesis, Univ. Aix-Marseille, France.

04/23: Reviewer of Yosra Marnissi PhD Thesis, Univ. Paris-Est, France.

10.2.3.2. Gaël Varoquaux

external reviewer, PhD defence of Christian Dansereau, Université de Montréal.

10.2.3.3. Alexandre Gramfort

12/19: Examiner of Thomas Moreau, Univ. Paris Saclay / ENS Paris-Saclay

12/13: Reviewer of Etienne Combrisson PhD Thesis, Univ. Lyon 1, France.

11/10: Reviewer of Lankinen Kaisu PhD Thesis, Univ. Aalto, Finland.

04/26: Reviewer of Seyed Mostafa Kia PhD Thesis, Trento FBK Univ., Italy.

10.2.3.4. Bertrand Thirion

18/01: Reviewer of Jonathan Vacher PhD thesis, Univ. Paris VI

3O/05: Examiner of Brahim Bougacha PhD thesis, University Nice

10.3. Popularization

10.3.1. Gaël Varoquaux

Présentation "Démocratisation de l'IA", Station F, (Dec 2017).

Présentation sur les logiciels libres de machine learning aux journées annuelles du groupe de travail sur les logiciels libres (Oct 2017)

Présentation sur scikit-learn au "Paris ML meetup" (Dec 2017)

Présentation sur scikit-learn au "Paris Pydata meetup" (Mars 2017)

10.3.2. Alexandre Gramfort

Présentation sur "L'intelligence artificielle au service de la société" à l'UNESCO dans le cadre des journées des ingénieurs et scientifiques de France (IESF) (Oct 2017). [0].

Présentation sur l'Intelligence Artificielle au conseil d'administration de l'Institut National du Cancer (INCA) (Nov 2017).

11. Bibliography

Major publications by the team in recent years

[1] D. Bzdok, M. Eickenberg, O. Grisel, B. Thirion, G. Varoquaux. *Semi-Supervised Factored Logistic Regression for High-Dimensional Neuroimaging Data*, in "Neural Information Processing Systems", December 2015, https://hal.archives-ouvertes.fr/hal-01211248.

[2] N. Chauffert, P. Ciuciu, J. Kahn, P. Weiss. *A Projection Method on Measures Sets*, in "Constructive Approximation", February 2017, vol. 45, n^{o} 1, p. 83 - 111 [*DOI :* 10.1007/s00365-016-9346-2], https://hal.inria.fr/hal-01432720.

[0] http://home.iesf.fr/752_p_49913/l-intelligence-artificielle.html

[3] M. EICKENBERG, E. DOHMATOB, B. THIRION, G. VAROQUAUX.*Total Variation meets Sparsity: statistical learning with segmenting penalties*, in "Medical Image Computing and Computer Aided Intervention (MICCAI)", München, Germany, Proceedings of MICCAI 2015, October 2015, https://hal.inria.fr/hal-01170619.

[4] A. A. HOYOS-IDROBO, G. VAROQUAUX, Y. SCHWARTZ, B. THIRION.*FReM – scalable and stable decoding with fast regularized ensemble of models*, in "NeuroImage", 2017, p. 1-16 [*DOI :* 10.1016/J.NEUROIMAGE.2017.10.005], https://hal.archives-ouvertes.fr/hal-01615015.

[5] M. JAS, D. A. ENGEMANN, Y. BEKHTI, F. A. RAIMONDO, A. GRAMFORT.*Autoreject: Automated artifact rejection for MEG and EEG data*, in "NeuroImage", 2017 [*DOI :* 10.1016/J.NEUROIMAGE.2017.06.030], https://hal.inria.fr/hal-01562403.

[6] A. MENSCH, J. MAIRAL, D. BZDOK, B. THIRION, G. VAROQUAUX.*Learning Neural Representations of Human Cognition across Many fMRI Studies*, in "Neural Information Processing Systems", Long Beach, United States, December 2017, https://hal.archives-ouvertes.fr/hal-01626823.

[7] A. MENSCH, J. MAIRAL, B. THIRION, G. VAROQUAUX.*Stochastic Subsampling for Factorizing Huge Matrices*, in "IEEE Transactions on Signal Processing", January 2018, vol. 66, n^{o} 1, p. 113-128 [*DOI :* 10.1109/TSP.2017.2752697], https://hal.archives-ouvertes.fr/hal-01431618.

[8] B. NG, G. VAROQUAUX, J.-B. POLINE, M. D. GREICIUS, B. THIRION.*Transport on Riemannian Manifold for Connectivity-based Brain Decoding*, in "IEEE Transactions on Medical Imaging", 2015, vol. PP, n^{o} 99, 9 [*DOI :* 10.1109/TMI.2015.2463723], https://hal.inria.fr/hal-01185200.

[9] F. PEDREGOSA, M. EICKENBERG, P. CIUCIU, B. THIRION, A. GRAMFORT.*Data-driven HRF estimation for encoding and decoding models*, in "NeuroImage", November 2014, p. 209–220, https://hal.inria.fr/hal-00952554.

[10] Y. SCHWARTZ, B. THIRION, G. VAROQUAUX.*Mapping cognitive ontologies to and from the brain*, in "NIPS (Neural Information Processing Systems)", United States, November 2013, http://hal.inria.fr/hal-00904763.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] E. DOHMATOB.*Enhancement of functional brain connectome analysis by the use of deformable models in the estimation of spatial decompositions of the brain images*, Université Paris-Saclay, September 2017, https://tel.archives-ouvertes.fr/tel-01630295.

Articles in International Peer-Reviewed Journal

[12] M. ALBUGHDADI, L. CHAARI, J.-Y. TOURNERET, F. FORBES, P. CIUCIU.*A Bayesian Non-Parametric Hidden Markov Random Model for Hemodynamic Brain Parcellation*, in "Signal Processing", June 2017, vol. 135, p. 132–146 [*DOI :* 10.1016/J.SIGPRO.2017.01.005], https://hal.archives-ouvertes.fr/hal-01426385.

[13] D. T. ALCALÁ-LÓPEZ, J. SMALLWOOD, E. JEFFERIES, F. B. VAN OVERWALLE, K. VOGELEY, R. B. MARS, B. B. TURETSKY, A. R. LAIRD, P. T. FOX, S. B. EICKHOFF, D. BZDOK.*Computing the Social Brain Connectome Across Systems and States*, in "Cerebral Cortex", 2017, p. 1 - 26 [*DOI :* 10.1093/CERCOR/BHX121], https://hal.archives-ouvertes.fr/hal-01519450.

[14] D. BZDOK.*Classical Statistics and Statistical Learning in Imaging Neuroscience: Two Statistical Cultures in Neuroimaging*, in "Frontiers in Human Neuroscience", November 2017, https://hal.archives-ouvertes.fr/hal-01583175.

[15] D. BZDOK, M. KRZYWINSKI, N. ALTMAN.*Machine learning: A primer*, in "Nature Methods", November 2017, https://hal.archives-ouvertes.fr/hal-01598285.

[16] D. BZDOK, M. KRZYWINSKI, N. ALTMAN.*Machine learning: Supervised methods, SVM and kNN*, in "Nature Methods", January 2018, p. 1-6, https://hal.archives-ouvertes.fr/hal-01657491.

[17] D. BZDOK, A. MEYER-LINDENBERG.*Machine learning for precision psychiatry: Opportunites and challenges*, in "Biological Psychiatry: Cognitive Neuroscience and Neuroimaging", February 2018, https://hal.archives-ouvertes.fr/hal-01643933.

[18] D. BZDOK, B. T. T. YEO.*Inference in the age of big data: Future perspectives on neuroscience*, in "NeuroImage", April 2017 [*DOI :* 10.1016/J.NEUROIMAGE.2017.04.061], https://hal.archives-ouvertes.fr/hal-01516891.

[19] N. CHAUFFERT, P. CIUCIU, J. KAHN, P. WEISS.*A Projection Method on Measures Sets*, in "Constructive Approximation", February 2017, vol. 45, n^{o} 1, p. 83 - 111 [*DOI :* 10.1007/S00365-016-9346-2], https://hal.inria.fr/hal-01432720.

[20] M. DUBOL, C. TRICHARD, A. LARISA SANDU, C. LEROY, M. RAHIM, J.-L. MARTINOT, E. ARTIGES.*286. Dopamine Transporter and Reward Anticipation in Psychiatric Patients: A Positron Emission Tomography and Functional Magnetic Resonance Imaging Study*, in "Biological Psychiatry", May 2017, vol. 81, n^{o} 10, p. S117 - S118 [*DOI :* 10.1016/J.BIOPSYCH.2017.02.300], https://hal.inria.fr/hal-01593040.

[21] M. DUBOL, C. TRICHARD, C. LEROY, A.-L. SANDU, M. RAHIM, B. GRANGER, E. T. TZAVARA, L. KARILA, J.-L. MARTINOT, E. ARTIGES.*Dopamine Transporter and Reward Anticipation in a Dimensional Perspective: A Multimodal Brain Imaging Study*, in "Neuropsychopharmacology", August 2017 [*DOI :* 10.1038/NPP.2017.183], https://hal.inria.fr/hal-01576551.

[22] T. DUPRÉ LA TOUR, L. TALLOT, L. GRABOT, V. DOYÈRE, V. VAN WASSENHOVE, Y. GRENIER, A. GRAMFORT.*Non-linear auto-regressive models for cross-frequency coupling in neural time series*, in "PLoS Computational Biology", December 2017, vol. 13, n^{o} 12, e1005893 [*DOI :* 10.1371/JOURNAL.PCBI.1005893], https://hal.archives-ouvertes.fr/hal-01679078.

[23] A. A. HOYOS-IDROBO, G. VAROQUAUX, Y. SCHWARTZ, B. THIRION.*FReM – scalable and stable decoding with fast regularized ensemble of models*, in "NeuroImage", 2017, p. 1-16 [*DOI :* 10.1016/J.NEUROIMAGE.2017.10.005], https://hal.archives-ouvertes.fr/hal-01615015.

[24] M. JAS, D. A. ENGEMANN, Y. BEKHTI, F. A. RAIMONDO, A. GRAMFORT.*Autoreject: Automated artifact rejection for MEG and EEG data*, in "NeuroImage", 2017, https://arxiv.org/abs/1612.08194 [*DOI :* 10.1016/J.NEUROIMAGE.2017.06.030], https://hal.inria.fr/hal-01562403.

[25] Y. LE GUEN, G. AUZIAS, F. LEROY, M. NOULHIANE, G. DEHAENE-LAMBERTZ, E. DUCHESNAY, J.-F. MANGIN, O. COULON, V. FROUIN.*Genetic Influence on the Sulcal Pits: On the Origin of the First Cortical Folds*, in "Cerebral Cortex", April 2017, p. 1 - 12 [*DOI :* 10.1093/CERCOR/BHX098], https://hal-amu.archives-ouvertes.fr/hal-01527005.

[26] J. LEFORT-BESNARD, D. S. BASSETT, J. SMALLWOOD, D. S. MARGULIES, B. DERNTL, O. GRUBER, A. ALEMAN, R. JARDRI, G. VAROQUAUX, B. THIRION, S. B. EICKHOFF, D. BZDOK.*Different shades of default mode disturbance in schizophrenia: Subnodal covariance estimation in structure and function*, in "Human Brain Mapping", January 2018, p. 1-52, https://hal.archives-ouvertes.fr/hal-01620441.

[27] G. LEMAITRE, F. NOGUEIRA, C. K. ARIDAS.*Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning*, in "Journal of Machine Learning Research", 2017, vol. 18, p. 1 - 5, https://hal.inria.fr/hal-01516244.

[28] J. LOULA, G. VAROQUAUX, B. THIRION.*Decoding fMRI activity in the time domain improves classification performance*, in "NeuroImage", August 2017 [*DOI :* 10.1016/J.NEUROIMAGE.2017.08.018], https://hal.inria.fr/hal-01576641.

[29] A. MENSCH, J. MAIRAL, B. THIRION, G. VAROQUAUX.*Stochastic Subsampling for Factorizing Huge Matrices*, in "IEEE Transactions on Signal Processing", January 2018, vol. 66, n^{o} 1, p. 113-128, https://arxiv.org/abs/1701.05363 [*DOI :* 10.1109/TSP.2017.2752697], https://hal.archives-ouvertes.fr/hal-01431618.

[30] B. NG, G. VAROQUAUX, J. BAPTISTE POLINE, B. THIRION, M. D. GREICIUS, K. L. POSTON.*Distinct alterations in Parkinson's Medication-state and Disease-state Connectivity Running Title: PD altered connectivity*, in "Neuroimage-Clinical", 2017, https://hal.archives-ouvertes.fr/hal-01614971.

[31] T. E. NICHOLS, S. DAS, S. B. EICKHOFF 3, A. C. EVANS, T. GLATARD, M. HANKE, N. KRIEGESKORTE, M. P. MILHAM, R. A. POLDRACK, J.-B. P. POLINE, E. P. PROAL, B. THIRION, D. VAN ESSEN, T. WHITE, B. T. YEO.*Best Practices in Data Analysis and Sharing in Neuroimaging using MRI*, in "Nature Neuroscience", February 2017, vol. 20, p. 299-303 [*DOI :* 10.1038/NN.4500], https://hal.inria.fr/hal-01577319.

[32] M. RAHIM, B. THIRION, D. BZDOK, I. BUVAT, G. VAROQUAUX.*Joint prediction of multiple scores captures better individual traits from brain images*, in "NeuroImage", June 2017 [*DOI :* 10.1016/J.NEUROIMAGE.2017.06.072], https://hal.inria.fr/hal-01547524.

[33] G. VAROQUAUX.*Cross-validation failure: small sample sizes lead to large error bars*, in "NeuroImage", June 2017, https://arxiv.org/abs/1706.07581 [*DOI :* 10.1016/J.NEUROIMAGE.2017.06.061], https://hal.inria.fr/hal-01545002.

[34] D. VATANSEVER, D. BZDOK, H.-T. WANG, G. MOLLO, M. SORMAZ, C. MURPHY, T. KARAPANAGIOTIDIS, J. SMALLWOOD, E. JEFFERIES.*Varieties of semantic cognition revealed through simultaneous decomposition of intrinsic brain connectivity and behaviour*, in "NeuroImage", July 2017, 34 [*DOI :* 10.1016/J.NEUROIMAGE.2017.06.067], https://hal.archives-ouvertes.fr/hal-01546394.

[35] H.-T. WANG, P. GUILIA, C. MURPHY, D. BZDOK, E. JEFFERIES, J. SMALLWOOD.*Dimensions of Experience: Exploring the Heterogeneity of the Wandering Mind*, in "Psychological Science", August 2017, https://hal.archives-ouvertes.fr/hal-01562681.

International Conferences with Proceedings

[36] J. ARIAS, P. CIUCIU, M. DOJAT, F. FORBES, A. FRAU-PASCUAL, T. PERRET, J. M. WARNKING.*PyHRF: A Python Library for the Analysis of fMRI Data Based on Local Estimation of the Hemodynamic Response Function*, in "16th Python in Science Conference (SciPy 2017)", Austin, TX, United States, July 2017 [*DOI :* 10.25080/SHINMA-7F4C6E7-006], https://hal.archives-ouvertes.fr/hal-01566457.

[37] E. B. Belilovsky, K. Kastner, G. Varoquaux, M. B. Blaschko.*Learning to Discover Sparse Graphical Models*, in "International Conference on Machine Learning", Sydney, Australia, August 2017, https://arxiv.org/abs/1605.06359 , https://hal.inria.fr/hal-01306491.

[38] P. Ciuciu, H. Wendt, S. Combrexelle, P. Abry.*Spatially regularized multifractal analysis for fMRI Data*, in "EMBC'17 - 39th International Conference of the IEEE Engineering in Medicine and Biology Society", Jeju, South Korea, Kwang Suk Park, July 2017, 4, https://hal.inria.fr/hal-01574216.

[39] M. Clerc, A. Gramfort, E. Olivi, T. Papadopoulo.*OpenMEEG software for forward problems handling non-nested geometries*, in "BaCI Conference 2017: International Conference on Basic and Clinical Multimodal Imaging", Bern, Switzerland, August 2017, https://hal.inria.fr/hal-01581710.

[40] A. Hoyos-Idrobo, G. Varoquaux, B. Thirion.*Towards a Faster Randomized Parcellation Based Inference*, in "PRNI 2017 - 7th International Workshop on Pattern Recognition in NeuroImaging", Toronto, Canada, June 2017, https://hal.inria.fr/hal-01552237.

[41] M. Jas, T. Dupré La Tour, U. Şimşekli, A. Gramfort.*Learning the Morphology of Brain Signals Using Alpha-Stable Convolutional Sparse Coding*, in "Advances in neural information processing systems", Long Beach, United States, December 2017, https://arxiv.org/abs/1705.08006 , https://hal.archives-ouvertes.fr/hal-01590988.

[42] C. Lazarus, P. Weiss, N. Chauffert, F. Mauconduit, M. Bottlaender, A. Vignaud, P. Ciuciu.*SPARKLING: Novel Non-Cartesian Sampling Schemes for Accelerated 2D Anatomical Imaging at 7T Using Compressed Sensing*, in "25th annua meeting of the International Society for Magnetic Resonance Imaging", Honolulu, United States, April 2017, https://hal.inria.fr/hal-01577200.

[43] C. Lazarus, P. Weiss, N. Chauffert, A. Vignaud, P. Ciuciu.*SPARKLING: nouveaux schémas d'échantillonnage compressif prospectif pour l'IRM haute résolution*, in "GRETSI", Juan les Pins, France, September 2017, https://hal.inria.fr/hal-01577207.

[44] G. Lemaitre, R. Martí, M. Rastgoo, F. Mériaudeau.*Computer-Aided Detection for Prostate Cancer Detection based on Multi-Parametric Magnetic Resonance Imaging*, in "EMBC 2017 : 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society", Jeju Island, South Korea, July 2017, https://hal.inria.fr/hal-01516245.

[45] A. Mensch, J. Mairal, D. Bzdok, B. Thirion, G. Varoquaux.*Learning Neural Representations of Human Cognition across Many fMRI Studies*, in "Neural Information Processing Systems", Long Beach, United States, December 2017, https://arxiv.org/abs/1710.11438 , https://hal.archives-ouvertes.fr/hal-01626823.

[46] M. Rahim, B. Thirion, G. Varoquaux.*Multi-output predictions from neuroimaging: assessing reduced-rank linear models*, in "PRNI 2017 - The 7th International Workshop on Pattern Recognition in Neuroimaging", Toronto, Canada, June 2017, p. 1 - 4 [*DOI :* 10.1109/PRNI.2017.7981504], https://hal.inria.fr/hal-01547572.

[47] M. Rahim, B. Thirion, G. Varoquaux.*Population-shrinkage of covariance to estimate better brain functional connectivity*, in "Medical Image Computing and Computer Assisted Interventions", Quebec city, Canada, September 2017, https://hal.inria.fr/hal-01547612.

Conferences without Proceedings

[48] D. BZDOK, M. EICKENBERG, G. VAROQUAUX, B. THIRION.*Hierarchical Region-Network Sparsity for High-Dimensional Inference in Brain Imaging*, in "International conference on Information Processing in Medical Imaging (IPMI) 2017", Boone, North Carolina, USA, June 2017, https://hal.archives-ouvertes.fr/hal-01480885.

[49] C. LAROCHE, H. PAPADOPOULOS, M. KOWALSKI, G. RICHARD.*Drum extraction in single channel audio signals using multi-layer non negative matrix factor deconvolution*, in "ICASSP", Nouvelle Orleans, United States, March 2017, https://hal.archives-ouvertes.fr/hal-01438851.

[50] M.-A. SCHULZ, G. VAROQUAUX, A. GRAMFORT, B. THIRION, D. BZDOK.*Label scarcity in biomedicine: Data-rich latent factor discovery enhances phenotype prediction*, in "Neural Information Processing Systems, Machine Learning in Health Workshop", Long Beach, United States, December 2017, https://hal.archives-ouvertes.fr/hal-01633096.

Other Publications

[51] P. ABLIN, J.-F. CARDOSO, A. GRAMFORT.*Faster ICA under orthogonal constraint*, November 2017, working paper or preprint, https://hal.inria.fr/hal-01651842.

[52] P. ABLIN, J.-F. CARDOSO, A. GRAMFORT.*Faster independent component analysis by preconditioning with Hessian approximations*, June 2017, https://arxiv.org/abs/1706.08171 - working paper or preprint, https://hal.inria.fr/hal-01552340.

[53] G. JULIA GUIOMAR NISO, K. J. GORGOLEWSKI, E. BOCK, T. BROOKS, G. FLANDIN, A. GRAMFORT, R. N. HENSON, M. JAS, V. LITVAK, J. MOREAU, R. OOSTENVELD, J.-M. SCHOFFELEN, F. TADEL, J. WEXLER, S. BAILLET.*MEG-BIDS: an extension to the Brain Imaging Data Structure for magnetoencephalography*, September 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01591080.

References in notes

[54] K. S. BUTTON, J. P. IOANNIDIS, C. MOKRYSZ, B. A. NOSEK, J. FLINT, E. S. ROBINSON, M. R. MUNAFÒ.*Power failure: why small sample size undermines the reliability of neuroscience*, in "Nature Reviews Neuroscience", 2013, vol. 14, n^{o} 5, p. 365–376.

[55] R. A. POLDRACK, C. I. BAKER, J. DURNEZ, K. J. GORGOLEWSKI, P. M. MATTHEWS, M. R. MUNAFÒ, T. E. NICHOLS, J.-B. POLINE, E. VUL, T. YARKONI.*Scanning the horizon: towards transparent and reproducible neuroimaging research*, in "Nature Reviews Neuroscience", 2017, vol. 18, n^{o} 2, p. 115–126.

Project-Team PARSIFAL

Proof search and reasoning with logic specifications

IN COLLABORATION WITH: Laboratoire d'informatique de l'école polytechnique (LIX)

IN PARTNERSHIP WITH:
CNRS

Ecole Polytechnique

RESEARCH CENTER
Saclay - Île-de-France

THEME
Proofs and Verification

Table of contents

Project-Team PARSIFAL

Creation of the Project-Team: 2007 July 01

Keywords:

Computer Science and Digital Science:

A2.1.1. - Semantics of programming languages
A2.1.11. - Proof languages
A2.4.2. - Model-checking
A2.4.3. - Proofs
A7.2. - Logic in Computer Science

Other Research Topics and Application Domains:

B9.4.1. - Computer science
B9.4.2. - Mathematics
B9.6. - Reproducibility

1. Personnel

Research Scientists

Dale Miller [Team leader, Inria, Senior Researcher]
Beniamino Accattoli [Inria, Researcher]
Kaustuv Chaudhuri [Inria, Researcher]
Francois Lamarche [Inria, Senior Researcher]
Stéphane Lengrand [CNRS, Researcher]
Gabriel Scherer [Inria, Researcher, from Sep 2017]
Lutz Straßburger [Inria, Researcher]

Technical Staff

Marco Volpe [Inria]

PhD Students

Roberto Blanco Martinez [Inria]
Andrea Condoluci [Bologna University, from Sep 2017]
Ulysse Gerard [Inria]
Quentin Heath [Inria, until Apr 2017]
Maico Carlos Leberle [Inria, from Oct 2017]
Matteo Manighetti [Inria, from May 2017]
Sonia Marin [Inria, until Oct 2017]

Post-Doctoral Fellow

Matteo Acclavio [Inria, from Oct 2017]

Visiting Scientists

Carlos Olarte [Universidade Federal do Rio Grande do Norte, Brazil, Jun 2017]
Elaine Pimentel [Universidade Federal do Rio Grande do Norte, Brazil, Jun 2017]

Administrative Assistant

Christine Aklouche [Inria]

2. Overall Objectives

2.1. Main themes

The aim of the Parsifal team is to develop and exploit *proof theory* and *type theory* in the specification, verification, and analysis of computational systems.

- *Expertise*: the team conducts basic research in proof theory and type theory. In particular, the team is developing results that help with automated deduction and with the manipulation and communication of formal proofs.
- *Design*: based on experience with computational systems and theoretical results, the team develops new logical principles, new proof systems, and new theorem proving environments.
- *Implementation*: the team builds prototype systems to help validate basic research results.
- *Examples*: the design and implementation efforts are guided by examples of specification and verification problems. These examples not only test the success of the tools but also drive investigations into new principles and new areas of proof theory and type theory.

The foundational work of the team focuses on *structural* and *analytic* proof theory, *i.e.*, the study of formal proofs as algebraic and combinatorial structures and the study of proof systems as deductive and computational formalisms. The main focus in recent years has been the study of the *sequent calculus* and of the *deep inference* formalisms.

An important research question is how to reason about computational specifications that are written in a *relational* style. To this end, the team has been developing new approaches to dealing with induction, co-induction, and generic quantification. A second important question is of *canonicity* in deductive systems, *i.e.*, when are two derivations "essentially the same"? This crucial question is important not only for proof search, because it gives an insight into the structure and an ability to manipulate the proof search space, but also for the communication of *proof objects* between different reasoning agents such as automated theorem provers and proof checkers.

Important application areas currently include:

- Meta-theoretic reasoning on functional programs, such as terms in the λ-calculus
- Reasoning about behaviors in systems with concurrency and communication, such as the π-calculus, game semantics, *etc.*
- Combining interactive and automated reasoning methods for induction and co-induction
- Verification of distributed, reactive, and real-time algorithms that are often specified using modal and temporal logics
- Representing proofs as documents that can be printed, communicated, and checked by a wide range of computational logic systems.
- Development of cost models for the evaluation of proofs and programs.

3. Research Program

3.1. General overview

There are two broad approaches for computational specifications. In the *computation as model* approach, computations are encoded as mathematical structures containing nodes, transitions, and state. Logic is used to *describe* these structures, that is, the computations are used as models for logical expressions. Intensional operators, such as the modals of temporal and dynamic logics or the triples of Hoare logic, are often employed to express propositions about the change in state.

The *computation as deduction* approach, in contrast, expresses computations logically, using formulas, terms, types, and proofs as computational elements. Unlike the model approach, general logical apparatus such as cut-elimination or automated deduction becomes directly applicable as tools for defining, analyzing, and animating computations. Indeed, we can identify two main aspects of logical specifications that have been very fruitful:

- *Proof normalization*, which treats the state of a computation as a proof term and computation as normalization of the proof terms. General reduction principles such as β-reduction or cut-elimination are merely particular forms of proof normalization. Functional programming is based on normalization [70], and normalization in different logics can justify the design of new and different functional programming languages [42].
- *Proof search*, which views the state of a computation as a a structured collection of formulas, known as a *sequent*, and proof search in a suitable sequent calculus as encoding the dynamics of the computation. Logic programming is based on proof search [75], and different proof search strategies can be used to justify the design of new and different logic programming languages [73].

While the distinction between these two aspects is somewhat informal, it helps to identify and classify different concerns that arise in computational semantics. For instance, confluence and termination of reductions are crucial considerations for normalization, while unification and strategies are important for search. A key challenge of computational logic is to find means of uniting or reorganizing these apparently disjoint concerns.

An important organizational principle is structural proof theory, that is, the study of proofs as syntactic, algebraic and combinatorial objects. Formal proofs often have equivalences in their syntactic representations, leading to an important research question about *canonicity* in proofs – when are two proofs “essentially the same?” The syntactic equivalences can be used to derive normal forms for proofs that illuminate not only the proofs of a given formula, but also its entire proof search space. The celebrated *focusing* theorem of Andreoli [44] identifies one such normal form for derivations in the sequent calculus that has many important consequences both for search and for computation. The combinatorial structure of proofs can be further explored with the use of *deep inference*; in particular, deep inference allows access to simple and manifestly correct cut-elimination procedures with precise complexity bounds.

Type theory is another important organizational principle, but most popular type systems are generally designed for either search or for normalization. To give some examples, the Coq system [84] that implements the Calculus of Inductive Constructions (CIC) is designed to facilitate the expression of computational features of proofs directly as executable functional programs, but general proof search techniques for Coq are rather primitive. In contrast, the Twelf system [79] that is based on the LF type theory (a subsystem of the CIC), is based on relational specifications in canonical form (*i.e.*, without redexes) for which there are sophisticated automated reasoning systems such as meta-theoretic analysis tools, logic programming engines, and inductive theorem provers. In recent years, there has been a push towards combining search and normalization in the same type-theoretic framework. The Beluga system [80], for example, is an extension of the LF type theory with a purely computational meta-framework where operations on inductively defined LF objects can be expressed as functional programs.

The Parsifal team investigates both the search and the normalization aspects of computational specifications using the concepts, results, and insights from proof theory and type theory.

3.2. Inductive and co-inductive reasoning

The team has spent a number of years in designing a strong new logic that can be used to reason (inductively and co-inductively) on syntactic expressions containing bindings. This work is based on earlier work by McDowell, Miller, and Tiu [72] [71] [76] [85], and on more recent work by Gacek, Miller, and Nadathur [4] [57]. The Parsifal team, along with our colleagues in Minneapolis, Canberra, Singapore, and Cachen, have been building two tools that exploit the novel features of this logic. These two systems are the following.

- Abella, which is an interactive theorem prover for the full logic.
- Bedwyr, which is a model checker for the “finite” part of the logic.

We have used these systems to provide formalize reasoning of a number of complex formal systems, ranging from programming languages to the λ-calculus and π-calculus.

Since 2014, the Abella system has been extended with a number of new features. A number of new significant examples have been implemented in Abella and an extensive tutorial for it has been written [1].

3.3. Developing a foundational approach to defining proof evidence

The team is developing a framework for defining the semantics of proof evidence. With this framework, implementers of theorem provers can output proof evidence in a format of their choice: they will only need to be able to formally define that evidence's semantics. With such semantics provided, proof checkers can then check alleged proofs for correctness. Thus, anyone who needs to trust proofs from various provers can put their energies into designing trustworthy checkers that can execute the semantic specification.

In order to provide our framework with the flexibility that this ambitious plan requires, we have based our design on the most recent advances within the theory of proofs. For a number of years, various team members have been contributing to the design and theory of *focused proof systems* [47] [50] [51] [52] [59] [68] [69] and we have adopted such proof systems as the corner stone for our framework.

We have also been working for a number of years on the implementation of computational logic systems, involving, for example, both unification and backtracking search. As a result, we are also building an early and reference implementation of our semantic definitions.

3.4. Deep inference

Deep inference [61], [63] is a novel methodology for presenting deductive systems. Unlike traditional formalisms like the sequent calculus, it allows rewriting of formulas deep inside arbitrary contexts. The new freedom for designing inference rules creates a richer proof theory. For example, for systems using deep inference, we have a greater variety of normal forms for proofs than in sequent calculus or natural deduction systems. Another advantage of deep inference systems is the close relationship to categorical proof theory. Due to the deep inference design one can directly read off the morphism from the derivations. There is no need for a counter-intuitive translation.

The following research problems are investigated by members of the Parsifal team:

- Find deep inference system for richer logics. This is necessary for making the proof theoretic results of deep inference accessible to applications as they are described in the previous sections of this report.
- Investigate the possibility of focusing proofs in deep inference. As described before, focusing is a way to reduce the non-determinism in proof search. However, it is well investigated only for the sequent calculus. In order to apply deep inference in proof search, we need to develop a theory of focusing for deep inference.

3.5. Proof nets, atomic flows, and combinatorial proofs

Proof nets graph-like presentations of sequent calculus proofs such that all "trivial rule permutations" are quotiented away. Ideally the notion of proof net should be independent from any syntactic formalism, but most notions of proof nets proposed in the past were formulated in terms of their relation to the sequent calculus. Consequently we could observe features like “boxes” and explicit “contraction links”. The latter appeared not only in Girard's proof nets [58] for linear logic but also in Robinson's proof nets [81] for classical logic. In this kind of proof nets every link in the net corresponds to a rule application in the sequent calculus.

Only recently, due to the rise of deep inference, new kinds of proof nets have been introduced that take the formula trees of the conclusions and add additional “flow-graph” information (see e.g., [67][5] leading to the notion of *atomic flow* and [62]. On one side, this gives new insights in the essence of proofs and their normalization. But on the other side, all the known correctness criteria are no longer available.

Combinatorial proofs [65] are another form syntax-independent proof presentation which separates the multiplicative from the additive behaviour of classical connectives.

The following research questions investigated by members of the Parsifal team:

- Finding (for classical and intuitionistic logic) a notion of canonical proof presentation that is deductive, i.e., can effectively be used for doing proof search.
- Studying the normalization of proofs using atomic flows and combinatorial proofs, as they simplify the normalization procedure for proofs in deep inference, and additionally allow to get new insights in the complexity of the normalization.
- Studying the size of proofs use combinatorial proofs.

3.6. Cost Models and Abstract Machines for Functional Programs

In the *proof normalization* approach, computation is usually reformulated as the evaluation of functional programs, expressed as terms in a variation over the λ-calculus. Thanks to its higher-order nature, this approach provides very concise and abstract specifications. Its strength is however also its weakness: the abstraction from physical machines is pushed to a level where it is no longer clear how to measure the complexity of an algorithm.

Models like Turing machines or RAM rely on atomic computational steps and thus admit quite obvious cost models for time and space. The λ-calculus instead relies on a single non-atomic operation, β-reduction, for which costs in terms of time and space are far from evident.

Nonetheless, it turns out that the number of β-steps is a reasonable time cost model, i.e.,it is polynomially related to those of Turing machines and RAM. For the special case of *weak evaluation* (i.e., reducing only β-steps that are not under abstractions)—which is used to model functional programming languages—this is a relatively old result due to Blelloch and Greiner [48] (1995). It is only very recently (2014) that the strong case—used in the implementation models of proof assistants—has been solved by Accattoli and Dal Lago [43].

With the recent recruitment of Accattoli, the team's research has expanded in this direction. The topics under investigations are:

1. *Complexity of Abstract Machines*. Bounding and comparing the overhead of different abstract machines for different evaluation schemas (weak/strong call-by-name/value/need λ-calculi) with respect to the cost model. The aim is the development of a complexity-aware theory of the implementation of functional programs.
2. *Reasonable Space Cost Models*. Essentially nothing is known about reasonable space cost models. It is known, however, that environment-based execution model—which are the mainstream technology for functional programs—do not provide an answer. We are exploring the use of the non-standard implementation models provided by Girard's Geometry of Interaction to address this question.

4. Application Domains

4.1. Integrating a model checker and a theorem prover

The goal of combining model checking with inductive and co-inductive theorem is appealing. The strengths of systems in these two different approaches are strikingly different. A model checker is capable of exploring a finite space automatically: such a tool can repeatedly explore all possible cases of a given computational space. On the other hand, a theorem prover might be able to prove abstract properties about a search space. For example, a model checker could attempt to discover whether or not there exists a winning strategy for, say, tic-tac-toe while an inductive theorem prover might be able to prove that if there is a winning strategy for one board then there is a winning strategy for any symmetric version of that board. Of course, the ability to combine proofs from these systems could drastically reduce the amount of state exploration and verification of proof certificates that are needed to prove the existence of winning strategies.

Our first step to providing an integration of model checking and (inductive) theorem proving was the development of a strong logic, that we call $\mathcal{G}$, which extends intuitionistic logic with notions of least and greatest fixed points. We had developed the proof theory of this logic in earlier papers [4] [57]. We have now recently converted the Bedwyr system so that it formally accepts almost all definitions and theorem statements that are accepted by the inductive theorem prover Abella. Thus, these two systems are proving theorems in the same logic and their results can now be shared.

Bedwyr's tabling mechanism has been extended so that its it can make use of previously proved lemmas. For instance, when trying to prove that some board position has a winning strategy, an available stored lemma can now be used to obtain the result if some symmetric board position is already in the table.

Heath and Miller have shown how model checking can be seen as constructing proof in (linear) logic [64]. For more about recent progress on providing checkable proof certificates for model checking, see the web site for Bedwyr http://slimmer.gforge.inria.fr/bedwyr/.

4.2. Implementing trusted proof checkers

Traditionally, theorem provers—whether interactive or automatic—are usually monolithic: if any part of a formal development was to be done in a particular theorem prover, then the whole of it would need to be done in that prover. Increasingly, however, formal systems are being developed to integrate the results returned from several, independent and high-performance, specialized provers: see, for example, the integration of Isabelle with an SMT solver [56] as well as the Why3 and ESC/Java systems.

Within the Parsifal team, we have been working on foundational aspects of this multi-prover integration problem. As we have described above, we have been developing a formal framework for defining the semantics of proof evidence. We have also been working on prototype checkers of proof evidence which are capable of executing such formal definitions. The proof definition language described in the papers [54], [53] is currently given an implementation in the λProlog programming language [74]. This initial implementation will be able to serve as a "reference" proof checker: others who are developing proof evidence definitions will be able to use this reference checker to make sure that they are getting their definitions to do what they expect.

Using λProlog as an implementation language has both good and bad points. The good points are that it is rather simple to confirm that the checker is, in fact, sound. The language also supports a rich set of abstracts which make it impossible to interfere with the code of the checker (no injection attacks are possible). On the negative side, the performance of our λProlog interpreters is lower than that of specially written checkers and kernels.

4.3. Trustworthy implementations of theorem proving techniques

Instead of integrating different provers by exchanging proof evidence and relying on a backend proof-checker, another approach to integration consists in re-implementing the theorem proving techniques as proof-search strategies, on an architecture that guarantees correctness.

Inference systems in general, and focused sequent calculi in particular, can serve as the basis of such an architecture, providing primitives for the exploration of the search space. These form a trusted *Application Programming Interface* that can be used to program and experiment various proof-search heuristics without worrying about correctness. No proof-checking is needed if one trusts the implementation of the API.

This approach has led to the development of the Psyche engine, and to its latest branch CDSAT.

Three major research directions are currently being explored, based on the above:

- The first one is about formulating automated reasoning techniques in terms of inference systems, so that they fit the approach described above. While this is rather standard for technique used in first-order Automated Theorem Provers (ATP), such as resolution, superposition, etc, this is much less standard in SMT-solving, the branch of automated reasoning that can natively handle reasoning in a combination of mathematical theories: the traditional techniques developed there usually organise the collaborations between different reasoning black boxes, whose opaque mechanisms less clearly connect to proof-theoretical inference systems. We are therefore investigating new foundations for reasoning in combinations of theories, expressed as fine-grained inference systems, and developed the *Conflict-Driven Satisfiability framework* for these foundations [19].

- The second one is about understanding how to deal with quantifiers in presence of one or more theories: On the one hand, traditional techniques for quantified problems, such as *unification* [41] or *quantifier elimination* are usually designed for either the empty theory or very specific theories. On the other hand, the industrial techniques for combining theories (Nelson-Oppen, Shostak, MCSAT [78], [82], [86], [66]) are designed for quantifier-free problems, and quantifiers there are dealt with incomplete *clause instantiation* methods or *trigger*-based techniques [55]. We are working on making the two approaches compatible.
- The above architecture's modular approach raises the question of how its different modules can safely cooperate (in terms of guaranteed correctness), while some of them are trusted and others are not. The issue is particularly acute if some of the techniques are run concurrently and exchange data at unpredictable times. For this we explore new solutions based on Milner's *LCF* [77]. In [60], we argued that our solutions in particular provide a way to fulfil the "Strategy Challenge for SMT-solving" set by De Moura and Passmore [87].

5. New Software and Platforms

5.1. Abella

FUNCTIONAL DESCRIPTION: Abella is an interactive theorem prover for reasoning about computations given as relational specifications. Abella is particuarly well suited for reasoning about binding constructs.

- Participants: Dale Miller, Gopalan Nadathur, Kaustuv Chaudhuri, Mary Southern, Matteo Cimini, Olivier Savary-Bélanger and Yuting Wang
- Partner: Department of Computer Science and Engineering, University of Minnesota
- Contact: Kaustuv Chaudhuri
- URL: http://abella-prover.org/

5.2. Bedwyr

Bedwyr - A proof search approach to model checking

FUNCTIONAL DESCRIPTION: Bedwyr is a generalization of logic programming that allows model checking directly on syntactic expressions that possibly contain bindings. This system, written in OCaml, is a direct implementation of two recent advances in the theory of proof search.

It is possible to capture both finite success and finite failure in a sequent calculus. Proof search in such a proof system can capture both may and must behavior in operational semantics. Higher-order abstract syntax is directly supported using term-level lambda-binders, the nabla quantifier, higher-order pattern unification, and explicit substitutions. These features allow reasoning directly on expressions containing bound variables.

The distributed system comes with several example applications, including the finite pi-calculus (operational semantics, bisimulation, trace analyses, and modal logics), the spi-calculus (operational semantics), value-passing CCS, the lambda-calculus, winning strategies for games, and various other model checking problems.

- Participants: Dale Miller, Quentin Heath and Roberto Blanco Martinez
- Contact: Quentin Heath
- URL: http://slimmer.gforge.inria.fr/bedwyr/

5.3. Checkers

Checkers - A proof verifier

KEYWORDS: Proof - Certification - Verification

FUNCTIONAL DESCRIPTION: Checkers is a tool in Lambda-prolog for the certification of proofs. Checkers consists of a kernel which is based on LKF and is based on the notion of ProofCert.

- Participants: Giselle Machado Nogueira Reis, Marco Volpe and Tomer Libal
- Contact: Tomer Libal
- URL: https://github.com/proofcert/checkers

5.4. Psyche

Proof-Search factorY for Collaborative HEuristics

FUNCTIONAL DESCRIPTION: Psyche is a modular platform for automated or interactive theorem proving, programmed in OCaml and built on an architecture (similar to LCF) where a trusted kernel interacts with plugins. The kernel offers an API of proof-search primitives, and plugins are programmed on top of the API to implement search strategies. This architecture is set up for pure logical reasoning as well as for theory-specific reasoning, for various theories.

RELEASE FUNCTIONAL DESCRIPTION: It is now equipped with the machinery to handle quantifiers and quantifier-handling techniques. Concretely, it uses meta-variables to delay the instantiation of existential variables, and constraints on meta-variables are propagated through the various branches of the search-space, in a way that allows local backtracking. The kernel, of about 800 l.o.c., is purely functional.

- Participants: Assia Mahboubi, Jean-Marc Notin and Stéphane Graham-Lengrand
- Contact: Stéphane Graham-Lengrand
- URL: http://www.lix.polytechnique.fr/~lengrand/Psyche/

6. New Results

6.1. Separating Functional Computation from Relations

Participants: Ulysse Gérard, Dale Miller.

The logical foundation of arithmetic generally starts with a quantificational logic over relations. Of course, one often wishes to have a formal treatment of functions within this setting. Both Hilbert and Church added choice operators (such as the epsilon operator) to logic in order to coerce relations that happen to encode functions into actual functions. Others have extended the term language with confluent term rewriting in order to encode functional computation as rewriting to a normal form (e.g., the Dedukti proof checking project [46]) It is possible to take a different approach that does not extend the underlying logic with either choice principles or with an equality theory. Instead, we use the familiar two-phase construction of focused proofs and capture functional computation entirely within one of these phases. As a result, computation of functions can remain purely relational even when it is computing functions. This result, which appearred in [22], could be used to add to the Abella theorem prover a primitive method for doing deterministic computations.

6.2. Translating between implicit and explicit versions of proof

Participants: Roberto Blanco, Zakaria Chihani, Dale Miller.

As we have demonstrated within the Parsifal team, the Foundational Proof Certificate (FPC) framework can be used to define the semantics of a wide range of proof evidence. We have given such definitions for a number of textbook proof systems as well as for the proof evidence output from some existing theorem proving systems. An important decision in designing a proof certificate format is the choice of how many details are to be placed within certificates. Formats with fewer details are smaller and easier for theorem provers to output but they require more sophistication from checkers since checking will involve some proof reconstruction. Conversely, certificate formats containing many details are larger but are checkable by less sophisticated checkers. Since the FPC framework is based on well-established proof theory principles, proof certificates can be manipulated in meaningful ways. In fact, we have shown how it is possible to automate moving from implicit to explicit (*elaboration*) and from explicit to implicit (*distillation*) proof evidence via the proof checking of a *pair of proof certificates*. Performing elaboration makes it possible to transform a proof certificate with details missing into a certificate packed with enough details so that a simple kernel (without support for proof reconstruction) can check the elaborated certificate. This design allows us to trust in only a single, simple checker of explicitly described proofs but trust in a range of theorem provers employing a range of proof structures. Experimental results of using this design appear in

6.3. Combinatorial Flows

Participant: Lutz Straßburger.

Combinatorial flows are a variation of combinatorial proofs that allow for the substitution of proofs into proofs (instead of just substituting formulas). This makes combinatorial flows p-equvalent to Frege systems with substitution, which are the strongest proof systems with respect to p-simulation, as studied in proof complexity. Since combinatorial flows have a polynomial correctness criterion, they can also be seen as an improvement to atomic flows (which do not have a correctness criterion). This work has been presented at the FCSD 2017 conference [37], [28]

6.4. Justification Logic for Constructive Modal Logic

Participants: Lutz Straßburger, Sonia Marin.

Justification logic is a family of modal logics generalizing the Logic of Proofs LP, introduced by Artemov in [45]. The original motivation, which was inspired by works of Kolmogorov and Gödel in the 1930's, was to give a classical provability semantics to intuitionistic propositional logic. The language of the Logic of Proofs can be seen as a modal language where occurrences of the $\Box$-modality are replaced with terms, also known as *proof polynomials*, *evidence terms*, or *justification terms*, depending on the setting. The intended meaning of the formula '$t : A$' is '*tis a proof ofA*' or, more generally, the reason for the validity of A. Thus, the justification language is viewed as a refinement of the modal language, with one provability construct $\Box$ replaced with an infinite family of specific proofs. In a joint work with Roman Kuznets (TU Wien), we add a second type of terms, which we call *witness terms* and denote by Greek letters. Thus, a formula $\diamond A$ is to be realized by '$\mu : A$'. The intuitive understanding of these terms is based on the view of $\diamond$ modality as representing consistency (with $\Box$ still read as provability). The term μ justifying the consistency of a formula is viewed as an abstract witnessing model for the formula. We keep these witnesses abstract so as not to rely on any specific semantics. All the operations on witness terms that we employ to ensure the realization theorem for CK, CD, CT, and $CS4$. This work has been presented at the IMLA 2017 workshop [40]

6.5. Proof Theory of Indexed Nested Sequents

Participants: Lutz Straßburger, Sonia Marin.

Indexed nested sequents are an extension of nested sequents allowing a richer underlying graph-structure that goes beyond the plain tree-structure of pure nested sequents. For this reason they can be used to give deductive systems to modal logics which cannot be captured by pure nested sequents. In this work we show how the standard cut-elimination procedure for nested sequents can be extended to indexed nested sequents, and we discuss how indexed nested sequents can be used for intuitionistic modal logics. These results have been presented at the TABLEAUX 2017 conference [24], [35]

6.6. On the Length of Medial-Switch-Mix Derivations

Participant: Lutz Straßburger.

Switch and medial are two inference rules that play a central role in many deep inference proof systems. In specific proof systems, the mix rule may also be present. In a joint work with Paola Bruscoli (University of Bath) we show that the maximal length of a derivation using only the inference rules for switch, medial, and mix, modulo associativity and commutativity of the two binary connectives involved, is quadratic in the size of the formula at the conclusion of the derivation. This shows, at the same time, the termination of the rewrite system. This result has been presented at the International Workshop on Logic, Language, Information, and Computation 2017 [20].

6.7. Maehara-style Modal Nested Calculi

Participant: Lutz Straßburger.

In a joint work with Roman Kuznets (TU Wien), we develop multi-conclusion nested sequent calculi for the fifteen logics of the intuitionistic modal cube between IK and IS5. The proof of cut-free completeness for all logics is provided both syntactically via a Maehara-style translation and semantically by constructing an infinite birelational countermodel from a failed proof search [83]. Interestingly, the Maehara-style translation for proving soundness syntactically fails due to the hierarchical structure of nested sequents. Consequently, we only provide the semantic proof of soundness. The countermodel construction used to prove completeness required a completely novel approach to deal with two independent sources of non-termination in the proof search present in the case of transitive and Euclidean logics.

6.8. Combining inference systems in the CDSAT framework

Participant: Stéphane Graham-Lengrand.

In 2016 we had designed a methodology [49], based on *inference systems*, for combining theories in SMT-solving, that supersedes the existing approaches, namely that of Nelson-Oppen [78] and that of MCSAT [86], [66]. While soundness and completeness of our approach were proved in 2016, we further developed, in 2017, the meta-theory of this system, now called CDSAT for *Conflict-Driven Satisfiability*, in particular with

- a proof of termination for the CDSAT system, and the identification of sufficient conditions, on the theory modules to be combined, for the global termination of the system to hold;
- a learning mechanism, whereby the system discovers lemmas along the run, which can be used later to speed-up the rest of the run;
- an enrichment of the CDSAT system with proof-object generation, and the identification of proof-construction primitives that can be used to make the answers produced by CDSAT correct-by-construction.

The first result, together with the introduction of the CDSAT framework, was publishing this year in [19]. The last two results are described in a paper accepted for publication at CPP in 2018.

6.9. Theory modules for CDSAT

Participant: Stéphane Graham-Lengrand.

The CDSAT system described above is a framework for the combination of theory modules, so it is only useful inasmuch many theories can be captured as CDSAT theory modules. Theory modules are essentially given by a set of inference rules and, for each input problem, a finite set of expressions that are allowed to be used by CDSAT at runtime. These ingredients need to satisfy some requirement for soundness, completess, and termination of CDSAT. In 2017 we identified such theory modules for the following theories

- Boolean logic;
- Linear Rational Arithmetic;
- Equality with Uninterpreted Function symbols;
- Any theory whose ground satisfiability is decidable, if one is willing to give up the fine-grained aspect of inference rules;
- Bitvectors (core fragment).

The first four cases of theories were published in [19], while the Bitvector theory was published in [21].

6.10. Environments and the Complexity of Abstract Machines

Participant: Beniamino Accattoli.

This joint work with Bruno Barras (Inria) [30] belongs to line of work *Cost Models and Abstract Machines for Functional Languages*, supported by the ANR project COCA HOLA.

We study various notions of environments (local, global, split) for abstract machines for functional languages, from a complexity and implementative point of view.

An environment is a data structure used to implement sharing of subterms. There are two main styles. The most common one is to have many local environments, one for every piece of code in the data structures of the machine. A minority of works uses a single global environment instead. Up to now, the two approaches have been considered equivalent, in particular at the level of the complexity of the overhead: they have both been used to obtain bilinear bounds, that is, linear in the number of beta steps and in the size of the initial term.

Our main result is that local environments admit implementations that are asymptotically faster than global environments, lowering the dependency from the size of the initial term from linear to logarithmic, thus improving the bounds in the literature. We also show that a third style, split environments, that are in between local and global ones, has the benefits of both. Finally, we provide a call-by-need machine with split environments for which we prove the new improved bounds on the overhead.

6.11. The Negligible and Yet Subtle Cost of Pattern Matching

Participant: Beniamino Accattoli.

This joint work with Bruno Barras (Inria) [31] belongs to line of work *Cost Models and Abstract Machines for Functional Languages*, supported by the ANR project COCA HOLA.

In this work we extend results about time cost models for the λ-calculus to a larger language, namely the λ-calculus with constructors and pattern matching. We consider all natural evaluation strategies, that is, call-by-name, call-by-value, and call-by-need.

The results are expected, and considered folklore, but we show that the question is subtler than it seems at first sight, by exhibiting some counter-example for naive formulations of the extensions. The, we show the actual results for the right extensions.

6.12. Implementing Open Call-by-Value

This joint work with Giulio Guerrieri (Oxford University) [32] belongs to line of work *Cost Models and Abstract Machines for Functional Languages*, supported by the ANR project COCA HOLA.

The theory of the call-by-value λ-calculus relies on weak evaluation and closed terms, that are natural hypotheses in the study of programming languages. To model proof assistants, however, strong evaluation and open terms are required. Open call-by-value is the intermediate setting of weak evaluation with open terms, on top of which Grégoire and Leroy designed the abstract machine of Coq. This paper provides a theory of abstract machines for open call-by-value. The literature contains machines that are either simple but inefficient, as they have an exponential overhead, or efficient but heavy, as they rely on a labelling of environments and a technical optimization. We introduce a machine that is simple and efficient: it does not use labels and it implements open call-by-value within a bilinear overhead. Moreover, we provide a new fine understanding of how different optimizations impact on the complexity of the overhead.

6.13. Further Formalizing the Meta-Theory of Linear Logic

Participants: Kaustuv Chaudhuri, Leonardo Lima, Giselle Reis.

We have continued our formalization of the meta-theory of substructural logics by giving a fully formal proof of cut-elimination (and hence of completeness) for focused classical first-order linear logic. This is the first time that this complete system has had a fully formalized proof.

This formalization serves as a *tour de force* of Abella's ability to reason about mutual induction and support sophisticated binding constructs.

An extended invited paper is currently under review, to possibly appear in a special issue of *Theoretical Computer Science* in 2018.

6.14. Formalized Meta-Theory of Simultaneous Substitutions

Participant: Kaustuv Chaudhuri.

It has long been claimed that a logical framework must have sophisticated built-in support for reasoning about formal substitutions in order to formalize relational meta-theorems such as strong normalization (using a logical relations style argument) or that applicative simulation is a pre-congruence. A number of type-theoretic frameworks in recent years, such as Beluga, have indeed started to incorporate such constructs in their core systems.

We have recently shown how to implement the meta-theory of simultaneous substitutions in the Abella system without any modification or extension of the (trusted) kernel, and without sacrificing any expressivity. The results of this paper will appear in the ACM Conference on Certified Programming in Jannuary 2018.

Our hope is that this work will be continued in the near future to build a specification language based on contextual LF in Abella, similar to how Abella/LF handles (ordinary) LF.

6.15. Hybrid Linear Logic Revisited

Participants: Kaustuv Chaudhuri, Joëlle Despeyroux, Carlos Olarte, Elaine Pimentel.

We have written a comprehensive account of hybrid linear logic (HyLL) and its relation to a number of related linear logic variants such as subexponential logic. One of the new and novel examples that we have fully worked out is how to encode CTL and CTL* in HyLL, which shows that HyLL can indeed serve as a logical framework for representing and reasoning about constrained transition systems, such as biochemical networks.

This account will appear in a special issue of MSCS in 2018.

6.16. Correctness of Speculative Optimizations with Dynamic Deoptimization

Participant: Gabriel Scherer.

This joint work with Olivier Flückiger, Ming-Ho Yee Ming-Ho, Aviral Goel, Amal Ahmed and Jan Vitek was initiated during Gabriel Scherer's post-doctoral stay at Northeastern University, Boston, USA.

Practitioners from the software industry find it difficult to implement Just-In-Time (JIT) compilers for dynamic programming languages, such as Javascript: they don't know how to reason on the correctness of their optimizations in the context of Just-In-Time code generation and deoptimization. We explain how to adapt reasoning approaches and proof techniques from standard compiler research to this new setting.

This work [14] will appear in POPL 2018.

7. Partnerships and Cooperations

7.1. European Initiatives

7.1.1. FISP: ANR blanc International

Participants: Kaustuv Chaudhuri, François Lamarche, Sonia Marin, Dale Miller, Lutz Straßburger.

Title: The Fine Structure of Formal Proof Systems and their Computational Interpretations

Duration: 01/01/2016 – 31/10/2019

Partners:

University Paris VII, PPS (PI: Michel Parigot)

Inria Saclay–IdF, EPI Parsifal (PI: Lutz Straßburger)

University of Innsbruck, Computational Logic Group (PI: Georg Moser)

Vienna University of Technology, Theory and Logic Group (PI: Matthias Baaz)

Total funding by the ANR: 316 805 EUR

The FISP project is part of an ambitious, long-term project whose objective is to apply the powerful and promising techniques from structural proof theory to central problems in computer science for which they have not been used before, especially the understanding of the computational content of proofs, the extraction of programs from proofs and the logical control of refined computational operations. So far, the work done in the area of computational interpretations of logical systems is mainly based on the seminal work of Gentzen, who in the mid-thirties introduced the sequent calculus and natural deduction, along with the cut-elimination procedure. But that approach shows its limits when it comes to computational interpretations of classical logic or the modelling of parallel computing. The aim of our project, based on the complementary skills of the teams, is to overcome these limits. For instance, deep inference provides new properties, namely full symmetry and atomicity, which were not available until recently and opened new possibilities at the computing level, in the era of parallel and distributed computing.

7.1.2. COCA HOLA: ANR JCJC Project

Participant: Beniamino Accattoli.

Title: COst model for Complexity Analyses of Higher-Order programming LAnguages.

Collaborators: Ugo Dal Lago (University of Bologna & Inria), Delia Kesner (Paris Diderot University), Damiano Mazza (CNRS & Paris 13 University), Claudio Sacerdoti Coen (University of Bologna).

Duration: 01/10/2016 – 31/09/2019

Total funding by the ANR: 155 280 EUR

The COCA HOLA project aims at developing complexity analyses of higher-order computations, i.e. that approach to computation where the inputs and outputs of a program are not simply numbers, strings, or compound data-types, but programs themselves. The focus is not on analysing fixed programs, but whole programming languages. The aim is the identification of adequate units of measurement for time and space, i.e. what are called reasonable cost models. The problem is non-trivial because the evaluation of higher-order languages is defined abstractly, via high-level operations, leaving the implementation unspecified. Concretely, the project will analyse different implementation schemes, measuring precisely their computational complexity with respect to the number of high-level operations, and eventually develop more efficient new ones. The goal is to obtain a complexity-aware theory of implementations of higher-order languages with both theoretical and practical downfalls.

The projects stems from recent advances on the theory of time cost models for the lambda-calculus, the computational model behind the higher-order approach, obtained by the principal investigator and his collaborators (who are included in the project).

COCA HOLA will span over three years and is organised around three work packages, essentially:

1. extending the current results to encompass realistic languages;
2. explore the gap between positive and negative results in the literature;
3. use ideas from linear logic to explore space cost models, about which almost nothing is known.

7.2. International Initiatives

7.2.1. Participation in Other International Programs

7.2.1.1. PHC Amadeus: Analytic Calculi for Modal Logics

Participants: Kaustuv Chaudhuri, Sonia Marin, Giselle Reis, Lutz Straßburger.

Title: Analytic Calculi for Modal Logics

Duration: 01/01/2016 – 31/12/2017

Austrian Partner: TU Wien, Institute for Computer Science (Department III)

Modal logics are obtained from propositional logics by adding modalities $\Box$ and $\diamond$, meaning necessity and possibility. Originally studied by philosophers in order to reason about knowledge and belief, modal logics have nowadays many applications in computer science. Well known examples are epistemic logics, which allow to formally reason about the knowledge of independently acting and interacting agents, temporal logics, which allow to reason about temporal properties of processes, and authentication logics, which are used to formally reason about authentication protocols.

The purpose of this project is to develop a proof theory for variants of modal logic that have applications in modern computer science but that have been neglected by traditional proof theory so far.

7.3. International Research Visitors

7.3.1. Visits of International Scientists

7.3.1.1. Internships

Riccardo Treglia was an intern funded by COCA HOLA during March, April, and May 2017. He was advised by Accattoli and worked on the complexity analysis of abstract machines for the λ-calculus.

7.3.2. Visits to International Teams

7.3.2.1. Research Stays Abroad

Stéphane Graham-Lengrand spent 8 months, from January 2017 to August 2017, at SRI International, Computer Science Lab. This visit developed a collaboration with N. Shankar, MP Bonacina, and D. Jovanovic, on new algorithms and new architectures for automated and interactive theorem proving, as well as on new programme verification techniques.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events Organisation

8.1.1.1. General Chair, Scientific Chair

D. Miller has been selected to be the LICS General Chair for three years starting in July 2018.

8.1.1.2. Member of the Organizing Committees

Lutz Straßburger was member of the organizing committee for the second FISP meeting in Paris

8.1.2. Scientific Events Selection

8.1.2.1. Chair of Conference Program Committees

D. Miller was the Program Committee chair for the FSCD'17: Second International Conference on Formal Structures for Computation and Deduction, Oxford, 3-6 September.

D. Miller was on the Steering Committee for the FSCD series of International Conference on Formal Structures for Computation and Deduction.

D. Miller was a member of the jury for selecting the 2017 Ackermann Award (the EACSL award for outstanding doctoral dissertation in the field of Logic in Computer Science).

D. Miller was a member of the 2012, 2016, and 2017 Herbrand Award Committee of the Association for Automated Reasoning.

D. Miller is also a member of the SIGLOG advisory board, starting November 2015.

B. Accattoli was one of the two Program Committee chairs of the 6th International Workshop on Confluence (IWC 2017).

K. Chaudhuri as a co-chair of the Program Committee for the workshop on Structures and Deduction, co-located with FSCD.

8.1.2.2. Member of the Conference Program Committees

D. Miller was on the Program Committee of the following international conferences.

- 26th International Conference on Automated Deduction, Gothenburg, Sweden, 6-11 August.

B. Accattoli was on the Program Committee of the following international workshops.

- LOLA 2017: Syntax and Semantics of Low-Level Languages, Reykjavik, Iceland, 19 June.
- WPTE 2017: 4th Workshop on Rewriting Techniques for Program Transformations and Evaluation, Oxford, UK, 8 September.
- DICE-FOPARA 2017: 8th Workshop on Developments in Implicit Computational complExity and 5th Workshop on Foundational and Practical Aspects of Resource Analysis, Uppsala, Sweden, 22–23 April.

S. Graham-Lengrand was on the Program Committee of the following international workshops.

- AFM 2017: Automated Formal Methods, Menlo Park, USA, 19 May.
- PxTP 2017: 5th Workshop on Proof eXchange for Theorem Proving, Brasilia, Brazil, 4 September.

K. Chaudhuri was on the Program Committee of the following international workshops.

- LFMTP 2017: Logical Frameworks and Meta-languages: Theory and Practice, Oxford, U.K.
- LSFA 2017: Logical and Semantic Frameworks with Applications, Brasilia, Brazil

G. Scherer was on the Program Committee of the following international conference.

- Trends in Functional Programming, University of Kent at Canterbury, UK, 19-21 June.

8.1.2.3. Reviewer

- Lutz Straßburger reviewed submissions for the following conferences: LICS 2017, LPAR-21, FoSSaCS 2018, LFCS 2018
- B. Accattoli was a reviewer for the international conferences LICS 2017 (twice) and FSCD 2017.
- F. Lamarche was reviewer for CSL 2017.
- S. Graham-Lengrand was a reviewer for the international conferences LICS 2017 (three times), CADE 2017, AFM 2017, CSL 2017, TYPES 2017, PxTP 2017, FOSSACS 2018.

- G. Scherer reviewed submissions for the following conferences: JFLA 2018, FoSSaCS 2018, as well as for the PriSC 2018 (Principle of Secure Compilation) workshop.

8.1.3. Journal

8.1.3.1. Member of the Editorial Boards

D. Miller is on the editorial board of the following journals: ACM Transactions on Computational Logic, Journal of Automated Reasoning (Springer), and Journal of Applied Logic (Elsevier).

K. Chaudhuri served as a guest editor for a special issue of Mathematical Structures of Computer Science devoted to Logical Frameworks.

8.1.3.2. Reviewer - Reviewing Activities

- Lutz Straßburger did reviewing work for the following journals: Journal of Applied Logic (JAL), Studia Logica, Mathematical Structures in Computer Science (MSCS), Logical Methods in Computer Science (LMCS), Journal of Logic and Computation (JLC), Journal of Automated Reasoning (JAR).
- B. Accattoli was a reviewer for the international journals Transactions on Computational Logic (TOCL, ACM), Mathematical Structures in Computer Science (MSCS, Cambridge University Press), Logical Methods in Computer Science (LMCS), Journal of Automated Reasoning (JAR, Springer), Annals of Pure and Applied Logic (APAL, Elsevier).
- S. Graham-Lengrand was a reviewer for the following international journals: Theory of Computing Systems (TOCS), Annals of Pure and Applied Logic (APAL), Mathematical Structures in Computer Science (MSCS), Logical Methods in Computer Science (LMCS), Journal of Automated Reasoning (JAR), Bulletin of Symbolic Logic (BSL).
- G. Scherer was a reviewer for the international journal Mathematical Structures in Computer Science (MSCS).

8.1.4. Invited Talks

- D. Miller gave invited talks at the following two regularly held international meetings.
  - LAP 2017: Sixth Conference on Logic and Applications, 18-22 September 2018, Dubrovnik, Croatia.
  - PADL 2017: Nineteenth International Symposium on Practical Aspects of Declarative Languages, 16-17 January 2017, Paris.
- Lutz Straßburger gave an invited talk at the 4th International Workshop on Structures and Deduction (SD 2017), affiliated with FSCD'17.
- B. Accattoli gave an invited talk at LSFA 2017, the 12th Workshop on Logical and Semantic Frameworks with Applications, Brasilia, Brazil, 23-24 September.
- S. Graham-Lengrand gave an invited talk at CSLI 2017, the 6th CSLI Workshop on Logic, Rationality & Intelligent Interaction, University of Stanford, Palo Alto, USA, 3-4 June.

8.1.5. Research Administration

L. Straßburger serves on the "commission développement technologique (CDT)" for Inria Saclay–Île-de-France (since June 2012).

F. Lamarche was "responsable de centre" Saclay – Ile de France for Raweb.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Master: D. Miller, "*MPRI 2-1: Logique linéaire et paradigmes logiques du calcul*", 12 hours, M2, Master Parisien de Recherche en Informatique, France.

Lutz Straßburger gave a course on “Efficient Proof Systems for Modal Logics” at ESSLLI 2017 (joint with Roman Kuznets, TU Wien)

Master: B. Accattoli, “*MPRI 2-1: Logique linéaire et paradigmes logiques du calcul*”, 9 hours, M2, Master Parisien de Recherche en Informatique, France.

B. Accattoli taught the mini-course *the complexity of β-reduction*, 3 hours, at the International School on Rewriting 2017, Eindhoven, The Netherlands, 3-7 July.

Licence: S. Graham-Lengrand, “*INF412: Fondements de l'Informatique: Logique, Modèles, Calcul*”, 32 hours eq. TD, L3, École Polytechnique, France.

Master: S. Graham-Lengrand, “*INF551: Computational Logic*”, 45 hours eq. TD, M1, École Polytechnique, France.

Licence: K. Chaudhuri, “*INF431*: Concurrence” and “*INF441*: Programmation avancée”, 80 hours eq. TD, L2, Ecole polytechnique, France.

8.2.2. Supervision

PhD in progress: Sonia Marin, 1 Nov 2014, supervised by L. Straßburger and D. Miller

PhD in progress: Roberto Blanco, Ulysse Gérard, and Matteo Manighetti, supervised by D. Miller

PhD in progress: François Thiré (since 1st October 2016), supervised by S. Graham-Lengrand (joint with G. Dowek)

8.2.3. Juries

D. Miller was a reporter for the habilitation of Olivier Hermant, 20 April 2017.

9. Bibliography

Major publications by the team in recent years

[1] D. BAELDE, K. CHAUDHURI, A. GACEK, D. MILLER, G. NADATHUR, A. TIU, Y. WANG.*Abella: A System for Reasoning about Relational Specifications*, in "Journal of Formalized Reasoning", 2014, vol. 7, n^{o} 2, p. 1-89 [*DOI :* 10.6092/ISSN.1972-5787/4650], https://hal.inria.fr/hal-01102709.

[2] K. CHAUDHURI, N. GUENOT, L. STRASSBURGER.*The Focused Calculus of Structures*, in "Computer Science Logic: 20th Annual Conference of the EACSL", Leibniz International Proceedings in Informatics (LIPIcs), Schloss Dagstuhl–Leibniz-Zentrum für Informatik, September 2011, p. 159–173, http://drops.dagstuhl.de/opus/frontdoor.php?source_opus=3229.

[3] M. FAROOQUE, S. GRAHAM-LENGRAND, A. MAHBOUBI.*A bisimulation between DPLL(T) and a proof-search strategy for the focused sequent calculus*, in "Proceedings of the 2013 International Workshop on Logical Frameworks and Meta-Languages: Theory and Practice (LFMTP 2013)", A. MOMIGLIANO, B. PIENTKA, R. POLLACK (editors), ACM Press, September 2013, http://dx.doi.org/10.1145/2503887.2503892.

[4] A. GACEK, D. MILLER, G. NADATHUR.*Nominal abstraction*, in "Information and Computation", 2011, vol. 209, n^{o} 1, p. 48–73, http://arxiv.org/abs/0908.1390.

[5] A. GUGLIELMI, T. GUNDERSEN, L. STRASSBURGER.*Breaking Paths in Atomic Flows for Classical Logic*, in "Proceedings of the 25th Annual IEEE Symposium on Logic in Computer Science (LICS 2010)", Edinburgh, United Kingdom, July 2010, p. 284–293 [*DOI :* 10.1109/LICS.2010.12], http://www.lix.polytechnique.fr/~lutz/papers/AFII.pdf.

[6] C. LIANG, D. MILLER.*Focusing and Polarization in Linear, Intuitionistic, and Classical Logics*, in "Theoretical Computer Science", 2009, vol. 410, n^{o} 46, p. 4747–4768.

[7] C. LIANG, D. MILLER.*A Focused Approach to Combining Logics*, in "Annals of Pure and Appl. Logic", 2011, vol. 162, n^{o} 9, p. 679–697 [*DOI :* 10.1016/J.APAL.2011.01.012], http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/lku.pdf.

[8] D. MILLER.*A proposal for broad spectrum proof certificates*, in "CPP: First International Conference on Certified Programs and Proofs", J.-P. JOUANNAUD, Z. SHAO (editors), LNCS, 2011, vol. 7086, p. 54–69, http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/cpp11.pdf.

[9] L. STRASSBURGER.*Extension without Cut*, in "Annals of Pure and Appl. Logic", 2012, vol. 163, n^{o} 12, p. 1995–2007.

[10] L. STRASSBURGER.*Combinatorial Flows and Their Normalisation*, in "2nd International Conference on Formal Structures for Computation and Deduction, FSCD 2017, September 3-9, 2017, Oxford, UK", D. MILLER (editor), LIPIcs, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2017, vol. 84, p. 31:1–31:17.

[11] A. TIU, D. MILLER.*Proof Search Specifications of Bisimulation and Modal Logics for the π-calculus*, in "ACM Trans. on Computational Logic", 2010, vol. 11, n^{o} 2, http://arxiv.org/abs/0805.2785.

Publications of the year

Articles in International Peer-Reviewed Journal

[12] B. ACCATTOLI, C. SACERDOTI COEN.*On the Value of Variables*, in "Information and Computation", August 2017, vol. 255, p. 224 - 242 [*DOI :* 10.1016/J.IC.2017.01.003], https://hal.archives-ouvertes.fr/hal-01675373.

[13] Z. CHIHANI, D. MILLER, F. RENAUD.*A Semantic Framework for Proof Evidence*, in "Journal of Automated Reasoning", 2017, vol. 59, n^{o} 3, p. 287-330 [*DOI :* 10.1007/S10817-016-9380-6], https://hal.inria.fr/hal-01390912.

[14] O. FLÜCKIGER, G. SCHERER, M.-H. YEE, A. GOEL, A. AHMED, J. VITEK. *Correctness of Speculative Optimizations with Dynamic Deoptimization*, in "Proceedings of the ACM on Programming Languages", 2017, https://arxiv.org/abs/1711.03050 [*DOI :* 10.1145/3158137], https://hal.inria.fr/hal-01646765.

[15] D. MILLER.*Proof Checking and Logic Programming*, in "Formal Aspects of Computing", 2017, vol. 29, n^{o} 3, p. 383-399 [*DOI :* 10.1007/S00165-016-0393-Z], https://hal.inria.fr/hal-01390901.

Articles in Non Peer-Reviewed Journal

[16] D. ILIK.*Perspectives for proof unwinding by programming languages techniques*, in "IfColog Journal of Logics and their Applications (FLAP)", November 2017, vol. 4, n^{o} 10, p. 3487-3508, https://arxiv.org/abs/1605.09177 , https://hal.inria.fr/hal-01354180.

International Conferences with Proceedings

[17] R. BLANCO, Z. CHIHANI, D. MILLER.*Translating Between Implicit and Explicit Versions of Proof*, in "CADE 26 - 26th International Conference on Automated Deduction", Gothenburg, Sweden, August 2017, https://hal.inria.fr/hal-01645016.

[18] R. BLANCO, D. MILLER, A. MOMIGLIANO.*Property-Based Testing via Proof Reconstruction Work-in-progress*, in "LFMTP 17: Logical Frameworks and Meta-Languages: Theory and Practice", Oxford, United Kingdom, September 2017, https://hal.inria.fr/hal-01646788.

[19] M. P. BONACINA, S. GRAHAM-LENGRAND, N. SHANKAR.*Satisfiability Modulo Theories and Assignments*, in "CADE 2017 - 26th International Conference on Automated Deduction", Gothenburg, Sweden, August 2017 [*DOI :* 10.1007/978-3-319-63046-5_4], https://hal.archives-ouvertes.fr/hal-01615830.

[20] P. BRUSCOLI, L. STRASSBURGER.*On the Length of Medial-Switch-Mix Derivations*, in "WoLLIC 2017 - Logic, Language, Information, and Computation", London, United Kingdom, July 2017, https://hal.inria.fr/hal-01635933.

[21] S. GRAHAM-LENGRAND, D. JOVANOVIĆ.*An MCSAT treatment of Bit-Vectors (preliminary report)*, in "SMT 2017 - 15th International Workshop on Satisfiability Modulo Theories", Heidelberg, Germany, July 2017, https://hal.archives-ouvertes.fr/hal-01615837.

[22] U. GÉRARD, D. MILLER.*Separating Functional Computation from Relations*, in "26th EACSL Annual Conference on Computer Science Logic (CSL 2017)", Stockholm, Sweden, LIPIcs, August 2017, vol. 82, n^{o} 23, p. 23:1–23:17, https://hal.inria.fr/hal-01615683.

[23] D. ILIK.*On the exp-log normal form of types: Decomposing extensional equality and representing terms compactly*, in "Proceedings of the 44th ACM SIGPLAN Symposium on Principles of Programming Languages", Paris, France, January 2017, p. 387-399, https://arxiv.org/abs/1502.04634 , https://hal.inria.fr/hal-01167162.

[24] S. MARIN, L. STRASSBURGER.*Proof theory for indexed nested sequents*, in "TABLEAUX 2017 - Automated Reasoning with Analytic Tableaux and Related Methods", Brasilia, Brazil, LNCS, September 2017, vol. 10501, p. 81-97, https://hal.inria.fr/hal-01635935.

[25] D. MILLER.*Linear logic as a logical framework*, in "Proceedings of Structures and Deduction (SD) 2017", Oxford, United Kingdom, September 2017, https://hal.archives-ouvertes.fr/hal-01615664.

[26] D. MILLER.*Mechanized Metatheory Revisited: An Extended Abstract* , in "Post-proceedings of TYPES 2016", Novi Sad, Serbia, 2017 [*DOI :* 10.4230/LIPICS], https://hal.inria.fr/hal-01615681.

[27] G. SCHERER.*Deciding equivalence with sums and the empty type*, in "POPL 2017", Paris, France, POPL 2017-Proceedings of the 44th ACM SIGPLAN Symposium on Principles of Programming Languages, January 2017, https://arxiv.org/abs/1610.01213 - This work was presented at POPL 2017: Principles of Programming Languages [*DOI :* 10.01213], https://hal.inria.fr/hal-01646064.

[28] L. STRASSBURGER.*Combinatorial Flows and their Normalisation*, in "FSCD 2017 - 2nd International Conference on Formal Structures for Computation and Deduction", Oxford, United Kingdom, Leibniz International Proceedings in Informatics (LIPIcs), September 2017, vol. 84, p. 311 - 3117 [*DOI :* 10.4230/LIPICS.FSCD.2017.31], https://hal.inria.fr/hal-01635931.

National Conferences with Proceeding

[29] G. BARANY, G. SCHERER.*Génération aléatoire de programmes guidée par la vivacité*, in "JFLA: Journées Francophones des Langages Applicatifs", Banyuls-sur-Mer, France, January 2018, https://hal.inria.fr/hal-01682691.

Conferences without Proceedings

[30] B. ACCATTOLI, B. BARRAS.*Environments and the Complexity of Abstract Machines*, in "The 19th International Symposium on Principles and Practice of Declarative Programming", Namur, Belgium, October 2017 [*DOI :* 10.1145/3131851.3131855], https://hal.archives-ouvertes.fr/hal-01675358.

[31] B. ACCATTOLI, B. BARRAS.*The Negligible and Yet Subtle Cost of Pattern Matching*, in "Programming Languages and Systems - 15th Asian Symposium", Suzhou, China, November 2017, https://hal.archives-ouvertes.fr/hal-01675369.

[32] B. ACCATTOLI, G. GUERRIERI.*Implementing Open Call-by-Value*, in "Fundamentals of Software Engineering - 7th International Conference", Teheran, Iran, April 2017, https://hal.archives-ouvertes.fr/hal-01675365.

[33] D. MILLER.*Using linear logic and proof theory to unify computational logic*, in "Proceedings of Trends in Linear Logic and Applications (TLLA 17)", Oxford, United Kingdom, September 2017, https://hal.archives-ouvertes.fr/hal-01615673.

Books or Proceedings Editing

[34] D. MILLER (editor). *2nd International Conference on Formal Structures for Computation and Deduction (FSCD 2017)*, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Oxford, United Kingdom, September 2017, https://hal.inria.fr/hal-01615598.

Research Reports

[35] S. MARIN, L. STRASSBURGER.*On the Proof Theory of Indexed Nested Sequents for Classical and Intuitionistic Modal Logics*, Inria, April 2017, n^{o} RR-9061, https://hal.inria.fr/hal-01515797.

[36] L. STRASSBURGER, R. KUZNETS.*Maehara-style Modal Nested Calculi*, Inria Saclay, November 2017, n^{o} RR-9123, https://hal.inria.fr/hal-01644750.

[37] L. STRASSBURGER.*Combinatorial Flows and Proof Compression*, Inria Saclay, March 2017, n^{o} RR-9048, https://hal.inria.fr/hal-01498468.

[38] L. STRASSBURGER.*Deep Inference, Expansion Trees, and Proof Graphs for Second Order Propositional Multiplicative Linear Logic*, Inria Saclay Ile de France, May 2017, n^{o} RR-9071, 38, https://hal.inria.fr/hal-01526831.

Other Publications

[39] J. COURTIEL, K. YEATS, N. ZEILBERGER.*Connected chord diagrams and bridgeless maps*, November 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01650141.

[40] R. KUZNETS, S. MARIN, L. STRASSBURGER.*Justification logic for constructive modal logic* *, July 2017, IMLA 2017 - 7th Workshop on Intuitionistic Modal Logic and Applications, https://hal.inria.fr/hal-01614707.

References in notes

[41] J. A. ROBINSON, A. VORONKOV (editors). *Handbook of Automated Reasoning*, Elsevier and MIT press, 2001.

[42] S. ABRAMSKY.*Computational Interpretations of Linear Logic*, in "Theoretical Computer Science", 1993, vol. 111, p. 3–57.

[43] B. ACCATTOLI, U. DAL LAGO.*Beta reduction is invariant, indeed*, in "Joint Meeting of the Twenty-Third EACSL Annual Conference on Computer Science Logic (CSL) and the Twenty-Ninth Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), CSL-LICS '14, Vienna, Austria, July 14 - 18, 2014", 2014, p. 8:1–8:10, http://doi.acm.org/10.1145/2603088.2603105.

[44] J.-M. ANDREOLI.*Logic Programming with Focusing Proofs in Linear Logic*, in "Journal of Logic and Computation", 1992, vol. 2, n^o 3, p. 297–347.

[45] S. N. ARTEMOV.*Operational modal logic*, Cornell University, 1995, n^o MSI 95–29.

[46] A. ASSAF, G. BUREL, R. CAUDERLIER, D. DELAHAYE, G. DOWEK, C. DUBOIS, F. GILBERT, P. HALMAGRAND, O. HERMANT, R. SAILLARD.*Dedukti: a Logical Framework based on the $\lambda\Pi$-Calculus Modulo Theory*, 2016, Unpublished, http://www.lsv.ens-cachan.fr/~dowek/Publi/expressing.pdf.

[47] D. BAELDE, D. MILLER, Z. SNOW.*Focused Inductive Theorem Proving*, in "Fifth International Joint Conference on Automated Reasoning (IJCAR 2010)", J. GIESL, R. HÄHNLE (editors), LNCS, 2010, n^o 6173, p. 278–292 [*DOI :* 10.1007/978-3-642-14203-1], http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/ijcar10.pdf.

[48] G. E. BLELLOCH, J. GREINER.*Parallelism in Sequential Functional Languages*, in "Proceedings of the seventh international conference on Functional programming languages and computer architecture, FPCA 1995, La Jolla, California, USA, June 25-28, 1995", 1995, p. 226–237, http://doi.acm.org/10.1145/224164.224210.

[49] M. P. BONACINA, S. GRAHAM-LENGRAND, N. SHANKAR.*A model-constructing framework for theory combination*, Universita degli Studi di Verona, September 2016, n^o RR-99/2016, https://hal.inria.fr/hal-01425305.

[50] K. CHAUDHURI.*The Focused Inverse Method for Linear Logic*, Carnegie Mellon University, December 2006, Technical report CMU-CS-06-162, http://reports-archive.adm.cs.cmu.edu/anon/2006/CMU-CS-06-162.pdf.

[51] K. CHAUDHURI, N. GUENOT, L. STRASSBURGER.*The Focused Calculus of Structures*, in "Computer Science Logic: 20th Annual Conference of the EACSL", Leibniz International Proceedings in Informatics (LIPIcs), Schloss Dagstuhl–Leibniz-Zentrum für Informatik, September 2011, p. 159–173 [*DOI :* 10.4230/LIPIcs.CSL.2011.159], http://drops.dagstuhl.de/opus/volltexte/2011/3229/pdf/16.pdf.

[52] K. CHAUDHURI, S. HETZL, D. MILLER.*A Multi-Focused Proof System Isomorphic to Expansion Proofs*, in "Journal of Logic and Computation", June 2014 [*DOI :* 10.1093/LOGCOM/EXU030], http://hal.inria.fr/hal-00937056.

[53] Z. CHIHANI, D. MILLER, F. RENAUD.*Checking Foundational Proof Certificates for First-Order Logic (extended abstract)*, in "Third International Workshop on Proof Exchange for Theorem Proving (PxTP 2013)", J. C. BLANCHETTE, J. URBAN (editors), EPiC Series, EasyChair, 2013, vol. 14, p. 58–66.

[54] Z. CHIHANI, D. MILLER, F. RENAUD.*Foundational proof certificates in first-order logic*, in "CADE 24: Conference on Automated Deduction 2013", M. P. BONACINA (editor), Lecture Notes in Artificial Intelligence, 2013, n° 7898, p. 162–177.

[55] C. DROSS, S. CONCHON, J. KANIG, A. PASKEVICH.*Reasoning with Triggers*, in "10th Intern. Worksh. on Satisfiability Modulo Theories, SMT 2012", P. FONTAINE, A. GOEL (editors), EPiC Series, EasyChair, June 2012, vol. 20, p. 22–31, http://www.easychair.org/publications/?page=2135488790.

[56] P. FONTAINE, J.-Y. MARION, S. MERZ, L. P. NIETO, A. TIU.*Expressiveness + Automation + Soundness: Towards Combining SMT Solvers and Interactive Proof Assistants*, in "TACAS: Tools and Algorithms for the Construction and Analysis of Systems, 12th International Conference", H. HERMANNS, J. PALSBERG (editors), LNCS, Springer, 2006, vol. 3920, p. 167–181, http://dx.doi.org/10.1007/11691372_11.

[57] A. GACEK, D. MILLER, G. NADATHUR.*Combining generic judgments with recursive definitions*, in "23th Symp. on Logic in Computer Science", F. PFENNING (editor), IEEE Computer Society Press, 2008, p. 33–44, http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/lics08a.pdf.

[58] J.-Y. GIRARD.*Linear Logic*, in "Theoretical Computer Science", 1987, vol. 50, p. 1–102.

[59] S. GRAHAM-LENGRAND, R. DYCKHOFF, J. MCKINNA.*A Focused Sequent Calculus Framework for Proof Search in Pure Type Systems*, in "Logical Methods in Computer Science", 2011, vol. 7, n° 1, http://www.lix.polytechnique.fr/~lengrand/Work/Reports/TTSC09.pdf.

[60] S. GRAHAM-LENGRAND.*Slot Machines: an approach to the Strategy Challenge in SMT solving (presentation only)*, in "13th International Workshop on Satisfiability Modulo Theories", San Francisco, United States, July 2015, https://hal.inria.fr/hal-01211209.

[61] A. GUGLIELMI.*A System of Interaction and Structure*, in "ACM Trans. on Computational Logic", 2007, vol. 8, n° 1.

[62] A. GUGLIELMI, T. GUNDERSEN.*Normalisation Control in Deep Inference Via Atomic Flows*, in "Logical Methods in Computer Science", 2008, vol. 4, n° 1:9, p. 1–36, http://arxiv.org/abs/0709.1205.

[63] A. GUGLIELMI, L. STRASSBURGER.*Non-commutativity and MELL in the Calculus of Structures*, in "Computer Science Logic, CSL 2001", L. FRIBOURG (editor), LNCS, Springer-Verlag, 2001, vol. 2142, p. 54–68.

[64] Q. HEATH, D. MILLER.*A framework for proof certificates in finite state exploration*, in "Proceedings of the Fourth Workshop on Proof eXchange for Theorem Proving", C. KALISZYK, A. PASKEVICH (editors), Electronic Proceedings in Theoretical Computer Science, Open Publishing Association, August 2015, n° 186, p. 11–26 [*DOI :* 10.4204/EPTCS.186.4], http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/pxtp2015.pdf.

[65] D. HUGHES.*Proofs Without Syntax*, in "Annals of Mathematics", 2006, vol. 164, n^{o} 3, p. 1065–1076.

[66] D. JOVANOVIĆ, C. BARRETT, L. DE MOURA.*The Design and Implementation of the Model Constructing Satisfiability Calculus*, in "Proc. of the 13th Int. Conf. on Formal Methods In Computer-Aided Design (FMCAD '13)", FMCAD Inc., 2013, Portland, Oregon, http://www.cs.nyu.edu/~barrett/pubs/JBdM13.pdf.

[67] F. LAMARCHE, L. STRASSBURGER.*Naming Proofs in Classical Propositional Logic*, in "Typed Lambda Calculi and Applications, TLCA 2005", P. URZYCZYN (editor), LNCS, Springer, 2005, vol. 3461, p. 246–261.

[68] C. LIANG, D. MILLER.*Focusing and Polarization in Linear, Intuitionistic, and Classical Logics*, in "Theoretical Computer Science", 2009, vol. 410, n^{o} 46, p. 4747–4768, http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/tcs09.pdf.

[69] C. LIANG, D. MILLER.*A Focused Approach to Combining Logics*, in "Annals of Pure and Applied Logic", 2011, vol. 162, n^{o} 9, p. 679–697, http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/lku.pdf.

[70] P. MARTIN-LÖF.*Constructive Mathematics and Computer Programming*, in "Sixth International Congress for Logic, Methodology, and Philosophy of Science", Amsterdam, North-Holland, 1982, p. 153–175.

[71] R. MCDOWELL, D. MILLER.*Reasoning with Higher-Order Abstract Syntax in a Logical Framework*, in "ACM Trans. on Computational Logic", 2002, vol. 3, n^{o} 1, p. 80–136, http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/mcdowell01.pdf.

[72] R. MCDOWELL, D. MILLER.*A Logic for Reasoning with Higher-Order Abstract Syntax*, in "Proceedings, Twelfth Annual IEEE Symposium on Logic in Computer Science", Warsaw, Poland, G. WINSKEL (editor), IEEE Computer Society Press, July 1997, p. 434–445.

[73] D. MILLER.*Forum: A Multiple-Conclusion Specification Logic*, in "Theoretical Computer Science", September 1996, vol. 165, n^{o} 1, p. 201–232.

[74] D. MILLER, G. NADATHUR.*Programming with Higher-Order Logic*, Cambridge University Press, June 2012, http://dx.doi.org/10.1017/CBO9781139021326.

[75] D. MILLER, G. NADATHUR, F. PFENNING, A. SCEDROV.*Uniform Proofs as a Foundation for Logic Programming*, in "Annals of Pure and Applied Logic", 1991, vol. 51, p. 125–157.

[76] D. MILLER, A. TIU.*A Proof Theory for Generic Judgments: An extended abstract*, in "Proc. 18th IEEE Symposium on Logic in Computer Science (LICS 2003)", IEEE, June 2003, p. 118–127, http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/lics03.pdf.

[77] R. MILNER.*LCF: A Way of Doing Proofs with a Machine*, in "Proc. of the 8th Intern. Symp. on the Mathematical Foundations of Computer Science", J. BECVÁR (editor), LNCS, Springer, 1979, vol. 74, p. 146-159.

[78] G. NELSON, D. C. OPPEN.*Simplification by Cooperating Decision Procedures*, in "ACM Press Trans. on Program. Lang. and Syst.", October 1979, vol. 1, n^{o} 2, p. 245–257, http://dx.doi.org/10.1145/357073.357079.

[79] F. PFENNING, C. SCHÜRMANN.*System Description: Twelf — A Meta-Logical Framework for Deductive Systems*, in "16th Conference on Automated Deduction", Trento, H. GANZINGER (editor), LNAI, Springer, 1999, n^{o} 1632, p. 202–206.

[80] B. PIENTKA, J. DUNFIELD.*Beluga: A Framework for Programming and Reasoning with Deductive Systems (System Description)*, in "Fifth International Joint Conference on Automated Reasoning", J. GIESL, R. HÄHNLE (editors), LNCS, 2010, n^{o} 6173, p. 15–21.

[81] E. P. ROBINSON.*Proof Nets for Classical Logic*, in "Journal of Logic and Computation", 2003, vol. 13, p. 777–797.

[82] R. E. SHOSTAK.*Deciding Combinations of Theories*, in "J. ACM", 1984, vol. 31, n^{o} 1, p. 1–12, http://dx.doi.org/10.1145/2422.322411.

[83] L. STRASSBURGER, R. KUZNETS.*Maehara-style Modal Nested Calculi*, Inria Saclay, November 2017, n^{o} RR-9123, https://hal.inria.fr/hal-01644750.

[84] THE COQ DEVELOPMENT TEAM.*The Coq Proof Assistant Version 8.3 Reference Manual*, Inria, October 2010.

[85] A. TIU, D. MILLER.*Proof Search Specifications of Bisimulation and Modal Logics for the π-calculus*, in "ACM Trans. on Computational Logic", 2010, vol. 11, n^{o} 2, http://arxiv.org/abs/0805.2785.

[86] L. M. DE MOURA, D. JOVANOVIC.*A Model-Constructing Satisfiability Calculus*, in "Proc. of the 14th Int. Conf. on Verification, Model Checking, and Abstract Interpretation (VMCAI'13)", R. GIACOBAZZI, J. BERDINE, I. MASTROENI (editors), LNCS, Springer-Verlag, 2013, vol. 7737, p. 1–12, http://dx.doi.org/10.1007/978-3-642-35873-9_1.

[87] L. M. DE MOURA, G. O. PASSMORE.*The Strategy Challenge in SMT Solving*, in "Automated Reasoning and Mathematics - Essays in Memory of William W. McCune", M. P. BONACINA, M. E. STICKEL (editors), LNCS, Springer, 2013, vol. 7788, p. 15–44 [*DOI :* 10.1007/978-3-642-36675-8_2], http://dx.doi.org/10.1007/978-3-642-36675-8.

Project-Team PETRUS

PErsonal & TRUSted cloud

IN PARTNERSHIP WITH:
Université Versailles Saint-Quentin

RESEARCH CENTER
Saclay - Île-de-France

THEME
Data and Knowledge Representation and Processing

Table of contents

Project-Team PETRUS

Creation of the Team: 2016 December 01, updated into Project-Team: 2017 July 01

Keywords:

Computer Science and Digital Science:
A1.1.8. - Security of architectures
A1.4. - Ubiquitous Systems
A3.1.2. - Data management, quering and storage
A3.1.3. - Distributed data
A3.1.5. - Control access, privacy
A3.1.6. - Query optimization
A3.1.8. - Big data (production, storage, transfer)
A3.1.9. - Database
A4.3. - Cryptography
A4.5. - Formal methods for security
A4.7. - Access control
A4.8. - Privacy-enhancing technologies

Other Research Topics and Application Domains:
B2.5.3. - Assistance for elderly
B6.4. - Internet of things
B6.6. - Embedded systems
B9.8. - Privacy

1. Personnel

Research Scientists
Nicolas Anciaux [Team leader, Inria, Researcher, HDR]
Luc Bouganim [Inria, Senior Researcher, HDR]

Faculty Members
Guillaume Scerri [Univ de Versailles Saint-Quentin-en-Yvelines, Associate Professor]
Philippe Pucheral [Univ de Versailles Saint-Quentin-en-Yvelines, Professor, HDR]
Iulian Sandu Popa [Univ de Versailles Saint-Quentin-en-Yvelines, Associate Professor]

External Collaborator
Benjamin Nguyen [INSA CVL, Professor, HDR]

Technical Staff
Aydogan Ersoz [Inria]
Oana Manea [Inria, until Oct 2017]
Laurent Schneider [Inria, from Aug 2017]

PhD Students
Paul Tran Van [Cozy Cloud (CIFRE), from 2014]
Axel Michel [Insa CVL, from 2015]
Riad Ladjel [Inria, from 2016]
Julien Loudet [Cozy Cloud (CIFRE), from 2016]
Dimitrios Tsolovos [Inria, from 2017]

Administrative Assistant
Emmanuelle Perrot [Inria]

2. Overall Objectives

2.1. Overall Objectives

We are witnessing an exponential accumulation of personal data on central servers: data automatically gathered by administrations and companies but also data produced by individuals themselves (e.g., photos, agendas, data produced by smart appliances and quantified-self devices) and deliberately stored in the cloud for convenience. The net effect is, on the one hand, an unprecedented threat on data privacy due to abusive usage and attacks and, on the other hand, difficulties in providing powerful user-centric services (e.g. personal big data) which require crossing data stored today in isolated silos. The Personal Cloud paradigm holds the promise of a Privacy-by-Design storage and computing platform, where each individual can gather her complete digital environment in one place and share it with applications and users, while preserving her control. However, this paradigm leaves the privacy and security issues in user's hands, which leads to a paradox if we consider the weaknesses of individuals' autonomy in terms of computer security, ability and willingness to administer sharing policies. The challenge is however paramount in a society where emerging economic models are all based - directly or indirectly - on exploiting personal data.

While many research works tackle the organization of the user's workspace, the semantic unification of personal information, the personal data analytics problems, the objective of the PETRUS project-team is to tackle the privacy and security challenges from an architectural point of view. More precisely, our objective is to help providing a technical solution to the personal cloud paradox. More precisely, our goals are (i) to propose new architectures (encompassing both software and hardware aspects) for secure personal cloud data management and formally prove important bricks of the architecture, (ii) propose new data administration models reaching the main requirements of a personal cloud (decentralized access and usage control models, data sharing, data collection and retention models, etc.) and study the enforcement of the resulting privacy policies based on secure hardware and formally proven architectural components, (iii) propose new secure distributed database indexing models, privacy preserving query processing strategies and data anonymization techniques for the personal cloud.

3. Research Program

3.1. Research program

To tackle the challenge introduced above, we identify four main lines of research:

- (Axis 1) Personal cloud server architectures. Based on the intuition that user control, security and privacy are key properties in the definition of trusted personal cloud solutions, our objective is to propose new architectures (encompassing both software and hardware aspects) for secure personal cloud data management and formally prove important bricks of the architecture.
- (Axis 2) Privacy preserving administration models and enforcement. This research axis is devoted to the definition of sharing rules that are easily manageable for the individual and enforced by default (i.e., secure implementation). Complementary to the definition of sharing policies, it is mandatory to help the average user regulate the complete lifecycle of her data, from its capture, to its dissemination and up to its deletion. Our objective is to propose new data administration models reaching the main requirements of a personal cloud (decentralized access and usage control models, data sharing, data collection and retention models, etc.) and study the enforcement of the resulting privacy policies based on secure hardware and formally proven architectural components.

- (Axis 3) Global query evaluation. The goal of this line of research is to provide capabilities for crossing data belonging to multiple individuals (e.g., performing statistical queries over personal data, computing queries on social graphs or organizing participatory data collection) in a fully decentralized setting while providing strong and personalized privacy guarantees. This means proposing new secure distributed database indexing models, privacy preserving query processing strategies and data anonymization techniques for the personal cloud.
- (Axis 4) Economic, legal and societal issues. This research axis is more transversal and entails multidisciplinary research, addressing the links between economic, legal, societal and technological aspects. We will follow here a multi-disciplinary approach based on a 3-step methodology: i) identifying important common issues related to privacy and to the exploitation of personal data; ii) characterizing their dimensions in all relevant disciplines and jointly study their entanglement; iii) validating the proposed analysis, models and trade-offs thanks to in vivo experiments.

These contributions will also rely on tools (algorithms, protocols, proofs, etc.) from other communities, namely security (cryptography, secure multiparty computations, formal methods, differential privacy, etc.) and distributed systems (distributed hash tables, gossip protocols, etc.). Beyond the research actions, we structure our software activity around a single common platform (rather than isolated demonstrators), integrating our main research contributions, called PlugDB. This platform is the cornerstone to help validating our research results through accurate performance measurements on a real platform, a common practice in the DB community, and target the best conferences. It is also a strong vector to federate the team, simplify the bootstrapping of new PhD or master students, conduct multi-disciplinary research and open the way to industrial collaborations and technological transfers.

4. Application Domains

4.1. Application Domains

As stated in the software section, the Petrus research strategy aims at materializing its scientific contributions in an advanced hardware/software platform with the expectation to produce a real societal impact. Hence, our software activity is structured around a common Secure Personal Cloud platform rather than several isolated demonstrators. This platform will serve as the foundation to develop a few emblematic applications. Several privacy-preserving applications can actually be targeted by a Personal Cloud platform, like: (i) smart disclosure applications allowing the individual to recover her personal data from external sources (e.g., bank, online shopping activity, insurance, etc.), integrate them and cross them to perform personal big data tasks (e.g., to improve her budget management) ; (ii) management of personal medical records for care coordination and well-being improvement; (iii) privacy-aware data management for the IoT (e.g., in sensors, quantified-self devices, smart meters); (iv) community-based sensing and community data sharing; (v) privacy-preserving studies (e.g., cohorts, public surveys, privacy-preserving data publishing). Such applications overlap with all the research axes described above but each of them also presents its own specificities. For instance, the smart disclosure applications will focus primarily on sharing models and enforcement, the IoT applications require to look with priority at the embedded data management and sustainability issues, while community-based sensing and privacy-preserving studies demand to study secure and efficient global query processing. Among these applications domains, one is already receiving a particular attention from our team. Indeed, we gained a strong expertise in the management and protection of healthcare data through our past DMSP (Dossier Medico-Social Partagé) experiment in the field. This expertise is being exploited to develop a dedicated healthcare and well-being personal cloud platform. We are currently deploying 10000 boxes equipped with PlugDB in the context of the DomYcile project. In this context, we are currently setting up an Inria Innovation Lab with the Hippocad company to industrialize this platform and deploy it at large scale (see Section the bilateral contract OwnCare II-Lab).

5. New Software and Platforms

5.1. PLUG-DB ENGINE

KEYWORDS: Databases - Personal information - Privacy - Hardware and Software Platform
FUNCTIONAL DESCRIPTION: PlugDB is a complete platform dedicated to a secure and ubiquitous management of personal data. It aims at providing an alternative to a systematic centralization of personal data. The PlugDB engine is a personal database server capable of storing data (tuples and documents) in tables and BLOBs, indexing them, querying them in SQL, sharing them through assertional access control policies and enforcing transactional properties (atomicity, integrity, durability). The PlugDB engine is embedded in a tamper-resistant hardware device combining the security of smartcard with the storage capacity of NAND Flash. The personal database is hosted encrypted in NAND Flash and the PlugDB engine code runs in the microcontroller. Complementary modules allow to pre-compile SQL queries for the applications, communicate with the DBMS from a remote Java program, synchronize local data with remote servers (typically used for recovering the database in the case of a broken or lost devices) and participate in distributed computation (e.g., global queries). PlugDB runs both on secure devices provided by Gemalto and on specific secure devices designed by PETRUS and assembled by electronic SMEs. Mastering the hardware platform opens up new research and experiment opportunities (e.g., we have recently integrated a Bluetooth module to communicate wirelessly with PlugDB and a fingerprint module to strongly authenticate users). PlugDB engine has been registered first at APP (Agence de Protection des Programmes) in 2009 - a new version being registered every two years and the hardware datasheets in 2015. PlugDB has been experimented in the field, notably in the healthcare domain. We also recently set up an educational platform on top of PlugDB, named SIPD (Système d'Information privacy-by-Design) and used at ENSIIE, INSA CVL and UVSQ through the Versailles Sciences Lab fablab, to raise students awareness of privacy protection problems and embedded programming. As a conclusion, PlugDB combines several research contributions from the team, at the crossroads of flash data management, embedded data processing and secure distributed computations. It then strongly federates all members of our team (permanent members, PhD students and engineers). It is also a vector of visibility, technological transfer and dissemination and gives us the opportunity to collaborate with researchers from other disciplines around a concrete privacy-enhancing platform. PlugDB is now being industrialized in the context of the OwnCare Inria Innovation Lab (II-Lab).

- Participants: Aydogan Ersoz, Laurent Schneider, Luc Bouganim, Nicolas Anciaux and Philippe Pucheral
- Contact: Nicolas Anciaux
- URL: https://project.inria.fr/plugdb/

6. New Results

6.1. Personal Cloud Architecture Based on Trusted Execution Environments (Axis 1)

Participants: Nicolas Anciaux [correspondent], Luc Bouganim, Riad Ladjel, Julien Loudet, Benjamin Nguyen, Philippe Pucheral, Iulian Sandu Popa, Guillaume Scerri, Paul Tran Van.

The Personal Cloud paradigm and its challenges: The time of individualized management and control over one's personal data is upon us. Thanks to smart disclosure initiatives, we can access our personal data from the companies or government agencies that collected them. Concurrently, Personal Cloud solutions are flourishing. Their goal is to empower us to leverage our personal data for our own good. However, managing our own personal data constitutes a considerable burden. We must now: (1) ensure the security of the data we gather; and (2) manage the disclosed data and control its usage. We inherit the combined responsibility of an information security expert and a database administrator. Since very few users are actually IT experts, personal cloud providers propose solutions to manage personal data on behalf of their customers. Thus, paradoxically, instead of empowering users, smart disclosure and personal clouds create new privacy risks. In this work, we formulate this paradox and the problems it creates. Our central contribution is a reference architecture for the Personal Cloud, instantiated on several hardware configuration using trusted execution environments (paper in preparation).

6.2. Data Management in Secure Hardware (Axis 1)

Participants: Nicolas Anciaux, Philippe Pucheral, Iulian Sandu Popa [correspondent].

Secure keyword search in the Personal Cloud: The Personal Cloud paradigm has emerged as a solution that allows individuals to manage under their control the collection, usage and sharing of their data. However, by regaining the full control over their data, the users also inherit the burden of protecting it against all forms of attacks and abusive usages. The Secure Personal Cloud architecture relieves the individual from this security task by employing a secure token (i.e., a tamper-resistant hardware device) to control all the sensitive information (e.g., encryption keys, metadata, indexes) and operations (e.g., authentication, data encryption/decryption, access control, and query processing). However, secure tokens are usually equipped with extremely low RAM but have significant Flash storage capacity (Gigabytes), which raises important barriers for embedded data management. This work [11] proposes a new embedded search engine specifically designed for secure tokens, which applies to the important use-case of managing and securing documents in the Personal Cloud context. Conventional search engines privilege either insertion or query scalability but cannot meet both requirements at the same time. Moreover, very few solutions support data deletions and updates in this context. In this work, we introduce three design principles, namely Write-Once Partitioning, Linear Pipelining and Background Linear Merging, and show how they can be combined to produce an embedded search engine matching the hardware constraints of secure tokens and reconciling high insert/delete/update rate and query scalability. Our experimental results, obtained with a prototype running on a representative hardware platform, demonstrate the scalability of the approach on large datasets and its superiority compared to state of the art methods. Finally, the integration of our solution in another important real use-case related to performing information retrieval in smart objects has been previously discussed in [5] and demonstrated at [25].

6.3. Data Management in Flash Memory (Axis 1)

Participant: Luc Bouganim [correspondent].

Understanding Flash I/O Pa erns on Open-Channel Solid-State Drives: Solid-State Drives (SSDs) have gained acceptance by providing the same block device abstraction as magnetic hard drives, at the cost of suboptimal resource utilization and unpredictable performance. Recently, Open-Channel SSDs have emerged as a means to obtain predictably high performance, based on a clean break from the block device abstraction. Open-channel SSDs embed a minimal flash translation layer (FTL) and expose their internals to the host. The Linux open-channel SSD subsystem, LightNVM, lets kernel modules as well as user-space applications control data placement and I/O scheduling. This way, it is the host that is responsible for SSD management. But what kind of performance model should the host rely on to guide the way it manages data placement and I/O scheduling? For addressing this question we have defined uFLIP- OC, a benchmark designed to identify the I/O patterns that are best suited for a given open-channel SSD. Our experiments on a Dragon- Fire Card (DFC) SSD, equipped with the OX controller, illustrate the performance impact of media characteristics and parallelism. In [17], we present uFLIP-OC and how it can be used to guide the design of host-based data systems on open-channel SSDs.

6.4. Data Sharing architecture for the Personal Cloud (Axis 2)

Participants: Nicolas Anciaux [correspondent], Philippe Pucheral, Paul Tran Van.

SWYSWYK Architecture: Pushed by recent legislation and smart disclosure initiatives, Personal Cloud platforms emerge and hold the promise of giving the control back to the individual on her data. However, this shift leaves the privacy and security issues in user's hands, a role that few people can properly endorse. Indeed, existing sharing models are difficult to administrate and securing their implementation in user's computing environment is an unresolved challenge. This study advocates the definition of a Privacy-by-Design sharing architecture, called SWYSWYK (Share What You See with Who You Know), dedicated to the Personal Cloud context. This architecture allows each user to physically visualize the net effects of sharing rules on her Personal Cloud and automatically provides tangible guarantees about the enforcement of the defined sharing policies. The architecture relies on a secure reference monitor, a set of user defined functions only interacting with the secure monitor and isolated from the unsecure environment, and an unsecure personal cloud platform managing encrypted personal data. The SWYSWYK architecture is presented in [20]. A validation of this architecture combining PlugDB to host the secure reference monitor, a RaspberryPI to launch the isolated user defined functions and a personal computer to host the untrusted personal cloud software was demonstrated in [19]. It shows the practicality of the approach and a performance evaluation on a real Personal Cloud platform.

6.5. Data sharing model for the Personal Cloud (Axis 2)

Participants: Nicolas Anciaux [correspondent], Paul Tran Van, Philippe Pucheral.

SWYSWYK Semantics: The personal cloud content intrinsically describes the individual's acquaintances under different forms (e.g., contact files, agendas, identity pictures, address book entries, etc.). Conversely, acquaintances are associated with pieces of information in the user's space (e.g., photos on which a friend appears). New sharing models should be thus able to map personal data to acquaintances (or subjects) and exploit their links with the stored documents (or objects) to produce authorizations satisfying users' sharing desires such as those expressed above. Interesting and common sharing rules could also be published and adopted by the members of a community of interest. In [18], we propose SWYSWYK, a new data sharing model which builds upon the transversal nature of the content of a personal cloud and makes easy and intuitive the definition and administration of sharing policies. Beyond the definition of the sharing policy, SWYSWYK provides means to the personal cloud owner to easily understand the net effects of a sharing policy, identify suspicious permissions and sanitize the sharing policy accordingly, and finally, to trust the way the policy is practically enforced. In [21] we demonstrate the semantics of the model with the goal to assess its practical interest for the personal cloud owner. To this end, we have integrated SWYSWYK in a real personal cloud platform (namely Cozy) and apply it to a smart surrounding scenario.

6.6. Privacy-preserving Computation Protocols on Asymmetric Architectures (Axis 3)

Participant: Iulian Sandu Popa [correspondent].

Distributed Vehicular Traffic Re-routing System for Congestion Avoidance: Centralized solutions for vehicular traffic re-routing to alleviate congestion suffer from two intrinsic problems: scalability, as the central server has to perform intensive computation and communication with the vehicles in real-time; and privacy, as the drivers have to share their location as well as the origins and destinations of their trips with the server. In this work [12], we proposed DIVERT, a distributed vehicular re-routing system for congestion avoidance. DIVERT offloads a large part of the re-routing computation at the vehicles, and thus, the re-routing process becomes practical in real-time. To take collaborative re-routing decisions, the vehicles exchange messages over vehicular ad hoc networks. DIVERT is a hybrid system because it still uses a server and Internet communication to determine an accurate global view of the traffic. In addition, DIVERT balances the user privacy with the re-routing effectiveness. The simulation results demonstrate that, compared with a centralized system, the proposed hybrid system increases the user privacy by 92 percent on average. In terms of average travel time, DIVERT's performance is slightly less than that of the centralized system, but it still achieves substantial gains compared to the no re-routing case. In addition, DIVERT reduces the CPU and network load on the server by 99.99 and 95 percent, respectively.

6.7. Privacy-preserving Anonymization Protocols on Asymmetric Architectures (Axis 3)

Participants: Axel Michel, Benjamin Nguyen [correspondent], Philippe Pucheral.

Managing Distributed Queries under Personalized Anonymity Constraints The benefit of performing Big data computations over individual's microdata is manifold, in the medical, energy or transportation fields to cite only a few, and this interest is growing with the emergence of smart disclosure initiatives around the world. However, these computations often expose microdata to privacy leakages, explaining the reluctance of individuals to participate in studies despite the privacy guarantees promised by statistical institutes. In this work [22], we propose a novel approach to push personalized privacy guarantees in the processing of database queries so that individuals can disclose different amounts of information (i.e. data at different levels of accuracy) depending on their own perception of the risk. Moreover, we propose a decentralized computing infrastructure based on secure hardware enforcing these personalized privacy guarantees all along the query execution process. A performance analysis conducted on a real platform shows the effectiveness of the approach.

6.8. Economic, legal and societal issues (Axis 4)

Participants: Nicolas Anciaux [correspondent], Philippe Pucheral.

Data Portability and Users' Empowerment as a Privacy Incentive. The principle of 'data portability' recently introduced in regulations (smart disclosure in the US, data portability in France and EU) is tightly coupled with the notion of Personal Cloud. We conduct a study of this principle in common with the DANTE Lab at UVSQ, in particular with Prof. Celia Zolynski (jurist, member of the CNN), within the DATAIA convergence institute at Inria and in the SIHS CNRS federation at UVSQ. Our recent contributions analyze the technical conditions under which individuals can get their data back from service providers according to this data portability principle, and examine its technical feasibility and legal opportunity. We also explain how data portability favors a form of users' empowerment, which can be viewed as a potential privacy incentive. Our recent results are presented in multi-disciplinary papers appeared in prestigious French journals like DALLOZ [14] and 'Revue Contrats, Concurrence, Consommation' [13] [15].

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

7.1.1. OwnCare II-Lab (Jul 2017 - Dec 2020)

Partners: PETRUS (Inria-UVSQ), Hippocad (SME)
Funding: to be determined
End 2016, the Yvelines district lauched a public call for tender to deploy an industrial solution aiming at covering the whole distinct (10.000 patients). The Hippocad company, in partnership with Inria, won this call for tender with a solution called DomYcile in May 2017 and the project was launched in July 2017. DomYcile is based on a home box combining the PlugDB hardware/software technology developed by the Petus team and a communication layer based on SigFox. Hippocad and Petrus then decided to launch a joint II-Lab (Inria Innovation Lab) named OwnCare. The objective is threefold: (1) build an industrial solution based on PlugDB and deploy it in the Yvelines district in the short-term, (2) use this Yvelines testbed to improve the solution and try to deploy it at the national/international level in the medium-term and (3) design flexible/secure/mobile personal medical folder solutions targeting individual uses rather than professional uses in the long-term. The DomYcile project with the Yvelynes district has started in July 2017 and the II-Lab should be officially created in January 2018.

7.2. Bilateral Grants with Industry

7.2.1. Cozy Cloud CIFRE - Tran Van contract (Oct 2014 -Feb 2017)

Partners: Cozy Cloud, PETRUS (Inria-UVSQ)
SMIS funding: 30k€
In relation with the bilateral contract mentioned above, a CIFRE PhD thesis has been started by Paul Tran Van. The objective is to capitalize on the Cozy-PlugDB platform to devise new access and usage control models to exchange data among devices of the same user (devices may have different levels of trustworthiness) and among different users thanks to a user-friendly sharing model (see the work on the SWYSWYK - Share What You See with Who You Know - model presented above).

7.2.2. Cozy Cloud CIFRE - Loudet contract (Apr 2016 - Apr 2019

Partners: Cozy Cloud, Inria-SMIS
SMIS funding: 45k€
In relation with the bilateral contract mentioned above, a second CIFRE PhD thesis has been started by Julien Loudet. The objective is to allow for a secure execution of distributed queries on a set of personal clouds associated to users, depending on social links, user's localization or user's profile. The general idea is to build secure indexes, distributed on the users' personal clouds and to devise a secure execution protocol revealing solely the query result to the querier. Such highly distributed secure queries potentially enable new (social) applications fed by user's personal data which could be developed on the Cozy-PlugDB platform.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR PerSoCloud (Jan 2017 - Jan 2020)

Partners: Orange Labs (coordinator), PETRUS (Inria-UVSQ), Cozy Cloud, U. of Versailles.
PETRUS funding: 170k€. The objective of PerSoCloud is to design, implement and validate a fullâ€‘fledged Privacy-by-Design Personal Cloud Sharing Platform. One of the major difficulties linked to the concept of personal cloud lies in organizing and enforcing the security of the data sharing while the data is no longer under the control of a central server. We identify three dimensions to this problem. Devices-sharing: assuming that the primary copy of user U1's personal data is hosted in a secure place, how to share and synchronize it with U1's multiple (mobile) devices without compromising security? Peers-sharing: how user U1 could exchange a subset of his-her data with an identified user U2 while providing to U1 tangible guarantees about the usage made by U2 of this data? Community-sharing: how user U1 could exchange a subset of his-her data with a large community of users and contribute to personal big data analytics while providing to U1 tangible guarantees about the preservation of his-her anonymity? In addition to tackling these three scientific and technical issues, a legal analysis will guarantee compliance of this platform with the security and privacy French and UE regulation, which firmly promotes the Privacy by Design principle, including the current reforms of personal data regulation.

8.1.2. PIA - PDP SECSi (May 2016 - Dec 2017)

Partners: Cozy Cloud (coordinator), Qwant, Inria (Inria-UVSQ), FING.
SMIS funding: 149k€. The objective of this PIA-PDP (Programme Investissement d'Avenir - Protection des DonnÃ©es Personnelles) SECSi project is to build a concrete Personal Cloud platform which can support a large scale deployment of Self Data services. Three major difficulties are identified and will be tackled in this project: (1) how to implement and enforce a fine control of the data flow when personal data are exploited by third party applications, (2) how to protect these same applications when processing is delegated to the personal cloud platform itself and (3) how to implement personalized search on the web without hurting user's privacy.

8.1.3. CityLab@Inria, Inria Project Lab (May 2014 -).

Inria Partners: ARLES-MIMOVE, CLIME, DICE, FUN, MYRIADS, OAK, PETRUS, URBANET, WILLOW. External partners: UC Berkeley.
Funding: not associated to individual project teams. CityLab@Inria studies ICT solutions toward smart cities that promote both social and environmental sustainability. A strong emphasis of the Lab is on the undertaking of a multi-disciplinary research program through the integration of relevant scientific and technology studies, from sensing up to analytics and advanced applications, so as to actually enact the foreseen smart city Systems of Systems. SMIS contributes to Privacy-by-Design architectures for trusted smart objects so as to ensure privacy to citizens, which is critical for ensuring that urbanscale sensing contributes to social sustainability and does not become a threat. The PhD Thesis of Dimitris Tsoulovos, co-directed by MIMOVE and PETRUS, is funded by CityLab. http://citylab.inria.fr/

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Selection

9.1.1.1. Member of the Conference Program Committees

- Nicolas Anciaux: VLDB'18, SIGMOD'17, PAP'17, DATA'17
- Luc Bouganim: Associate Editor for VLDB'18, VLDB'17, BDA'17
- Philippe Pucheral: EDBT'17, DATA'17, MOBILITY'17
- Iulian Sandu Popa: APWeb-WAIM'17, DATA'17, IEEE MobileCloud'17

9.1.1.2. Reviewer

- Guillaume Scerri: CCS'17

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards

- Nicolas Anciaux: Associate Editor of the VLDB Journal

9.1.2.2. Reviewer - Reviewing Activities

- Iulian Sandu Popa: ACM Transactions on Spatial Algorithms and Systems, International Journal of Geo-Information, Journal of Intelligent Transportation Systems

9.1.3. Invited Talks

- Nicolas Anciaux: "A new Approach for the Secure Personal Cloud", Security Seminar, Loria, 23 Mar. 2017. http://seminaire-securite.loria.fr/seances-passees.fr.html

9.1.4. Research Administration

Philippe Pucheral: Member of the HDR committee of the STV doctoral school (UVSQ) since 2014

Philippe Pucheral: Member of the steering committee of the ED STIC doctoral school of University Paris-Saclay, 'Data, Knowledge and Interactions' committee (about 250 PhD students) since 2014

Nicolas Anciaux: Member of the council of the doctoral college of the University Paris-Saclay since 2017

Nicolas Anciaux: Correspondent for Inria Saclay at ED STIC doctoral school of University Paris-Saclay since 2017

Nicolas Anciaux: Responsible for the 'Mission Jeunes Chercheurs' at Inria Saclay since 2017

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence : Iulian Sandu Popa, Bases de données (niveau L3), 96, UVSQ, France. Guillaume Scerri , Initiation aux bases de données (niveau L2), 63, UVSQ, France. Guillaume Scerri, Fondements de l'informatique (niveau L1), 72, UVSQ, France. Guillaume Scerri, Théorie des Langages (niveau L2), 36, UVSQ, France.

Master : Iulian Sandu Popa, Bases de données relationnelles (niveau M1), Gestion des données spatiotemporelles (niveau M2), Sécurité des bases de données (niveau M2), 96, UVSQ, France. Philippe Pucheral, responsible of the DataScale master, courses in M1 and M2 in databases and in security, introductory courses for jurists,UVSQ, France. Luc Bouganim, Bases de données relationnelles et XML (niveau M1 et M2), 20, AFTI, France. Guillaume Scerri, Bases de données relationnelles (niveau M1), 36, UVSQ, France. Guillaume Scerri, Sécurité et bases de données pour juristes, 4.5, UVSQ, France.

Engineers school : Nicolas Anciaux, courses on Databases (module IN206, niveau M1), 21, and Advanced databases (module ASI13, niveau M2), 24, at ENSTA ParisTech. Nicolas Anciaux, Systèmes d'Information "privacy by design" (niveau M1), 15, at ENSIIE Evry, France. Luc Bouganim, Systèmes d'Information "privacy by design" (niveau M1), 38, ENSIIE Evry et INSA CVL, France. Luc Bouganim, Bases de données et sécurité (niveau M2), 20, Telecom ParisTech.

9.2.2. Supervision

PhD in progress : Paul Tran Van, Partage de documents sécurisé dans le Cloud Personnel, October 2014, Nicolas Anciaux and Philippe Pucheral

PhD in progress : Axel Michel, Secure Distributed Computations, October 2015, Benjamin Nguyen and Philippe Pucheral

PhD in progress : Julien Loudet, Personal Queries on Personal Clouds, July 2016, Luc Bouganim and Iulian Sandu Popa

PhD in progress : Riad Ladjel, Secure Distributed Computation for the Personal Cloud, October 2016, Nicolas Anciaux and Philippe Pucheral

PhD in progress : Dimitri Tsoulovos, Privacy-by-design Middleware for Urban-scale Mobile Crowdsensing, April 2017, Nicolas Anciaux and Valérie Issarny (Inria Mimove)

9.2.3. Juries

Philippe Pucheral : reviewer of the PhD thesis of Yifan Li (CNAM Paris, 15/12/2017)

Nicolas Anciaux : Jury member of the PhD thesis of David Montoya (ENS Paris-Saclay, 6/3/2017)

Luc Bouganim : Reviewer of the HDR of Jalil Boukhobza (University of Bretagne Sud, 4/7/2017)

Luc Bouganim : Reviewer of the PhD of Levent Demir (University of Grenobles Alpes, 7/12/2017)

9.3. Popularization

MOOC "Défis technologiques des villes intelligentes participatives", du 6 Nov. 2017 au 31 Dec. 2017. Auteur(s): Nicolas Anciaux, Stéphane Grumbach, Valérie Issarny, Nathalie Mitton, Christine Morin, Animesh Pathak, Hervé Rivano. http://www.sup-numerique.gouv.fr/cid94187/mooc-villes-intelligentes-defis-technologiques-et-societaux.html

EIT Digital Professional School course "Technological challenges of participatory Smart Cities". Course for companies and cities. By Nicolas Anciaux, Stéphane Grumbach, Valérie Issarny, Nathalie Mitton, Christine Morin, Animesh Pathak, Hervé Rivano. https://www.eitdigital.eu/newsroom/news/article/eit-digital-professional-school-launches-a-smart-cities-course-for-companies-and-cities/

Nicolas Anciaux: interview pour le magazine La Recherche. Numéro de Juin 2017, dossier "Données personnelles: Notre vie privée est-elle en danger?", pp.81-86, par Vincent Glavieux et Denis Delbecq.

Rencontre Inria - Industrie, Les données et leurs applications, Conférence "Sécurité des données", Philippe Pucheral et Eric Léandri (fondateur et PDG de Qwant), 6 Juin 2017. https://www.inria.fr/centre/paris/innovation/rii-les-donnees-17-18-octobre-2017/programme

Rencontre Inria - Industrie, Les données et leurs applications, Démonstration du projet Domycile, Nicolas Anciaux, , 6 Juin 2017. https://www.inria.fr/centre/paris/innovation/rii-les-donnees-17-18-octobre-2017/demos/domycile

Rencontre Inria - Industrie, Les données et leurs applications, Démonstration du projet Domycile, Paul Tran Van, , 6 Juin 2017.

Round table: Accès aux données et aux compétences, Convention Systematic Paris-Région, Luc Bouganim, June 2017.

Demonstration: PlugDB and the Secure Mobile Laboratory, Turing building inauguration, Luc Bouganim, Iulian Sandu Popa, Riad Ladjel, February 2017.

10. Bibliography

Major publications by the team in recent years

[1] T. Allard, B. Nguyen, P. Pucheral.*MetaP: Revisiting Privacy-Preserving Data Publishing using Secure Devices*, in "Distributed and Parallel Databases", June 2014, vol. 32, n° 1, p. 191-244 [*DOI :* 10.1007/s10619-013-7122-x], https://hal.archives-ouvertes.fr/hal-00934586.

[2] N. Anciaux, P. Bonnet, L. Bouganim, B. Nguyen, P. Pucheral, I. Sandu Popa.*Trusted Cells : A Sea Change for Personnal Data Services*, in "CIDR 2013 - 6th Biennal Conference on Innovative Database Research", Asilomar, United States, 2013, 4, http://hal.inria.fr/hal-00768379.

[3] N. Anciaux, L. Bouganim, T. Delot, S. Ilarri, L. Kloul, N. Mitton, P. Pucheral.*Folk-IS: Opportunistic Data Services in Least Developed Countries*, in "40th International Conference on Very Large Data Bases (VLDB)", Hangzhou, China, Zhejiang University, September 2014, https://hal.inria.fr/hal-00906204.

[4] N. Anciaux, L. Bouganim, P. Pucheral, Y. Guo, L. Le Folgoc, S. Yin.*MILo-DB: a personal, secure and portable database machine*, in "Distributed and Parallel Databases", March 2014, vol. 32, n° 1, p. 37-63 [*DOI :* 10.1007/s10619-012-7119-x], https://hal.archives-ouvertes.fr/hal-00768355.

[5] N. Anciaux, S. Lallali, I. Sandu Popa, P. Pucheral.*A Scalable Search Engine for Mass Storage Smart Objects*, in "Proceedings of the 41th International Conference on Very Large Databases (VLDB)", Kohala Coast, Hawaii, United States, August 2015, vol. 8, n° 9, p. 910-921 [*DOI :* 10.14778/2777598.2777600], https://hal.inria.fr/hal-01176458.

[6] N. Anciaux, B. Nguyen, I. Sandu Popa.*Tutorial: Managing Personal Data with Strong Privacy Guarantees*, March 2014, p. 672-673 [*DOI :* 10.5441/002/edbt.2014.71], https://hal.inria.fr/hal-01096633.

[7] G. Scerri, B. Warinschi, M. Barbosa, B. Portela.*Foundations of Hardware-Based Attested Computation and Application to SGX*, March 2016, p. 245-260 [*DOI :* 10.1109/EuroSP.2016.28], https://hal.inria.fr/hal-01417137.

[8] C. Q. TO, B. NGUYEN, P. PUCHERAL.*TrustedMR: A Trusted MapReduce System based on Tamper Resistance Hardware* , in "Proceedings of the 23rd International Conference on Cooperative Information Systems (COOPIS)", Rhodes, Greece, October 2015, p. 38-56 [*DOI :* 10.1007/978-3-319-26148-5_3], https://hal.inria.fr/hal-01254951.

[9] C. Q. TO, B. NGUYEN, P. PUCHERAL.*Private and Scalable Execution of SQL Aggregates on a Secure Decentralized Architecture*, in "ACM Transactions on Database Systems", 2016, vol. 41, n^{o} 3, p. 16:1-16:43, https://hal.archives-ouvertes.fr/hal-01296432.

[10] D. H. TON THAT, I. SANDU POPA, K. ZEITOUNI.*TRIFL: A Generic Trajectory Index for Flash Storage*, in "ACM Transactions on Algorithms", July 2015, vol. 1, n^{o} 2, 44 [*DOI :* 10.1145/2786758], https://hal.inria.fr/hal-01176563.

Publications of the year

Articles in International Peer-Reviewed Journal

[11] S. LALLALI, N. ANCIAUX, I. SANDU POPA, P. PUCHERAL.*Supporting secure keyword search in the personal cloud*, in "Information Systems", December 2017, vol. 72, p. 1 - 26 [*DOI :* 10.1016/J.IS.2017.09.003], https://hal.inria.fr/hal-01660599.

[12] S. J. PAN, I. SANDU POPA, C. BORCEA.*DIVERT: A Distributed Vehicular Traffic Re-Routing System for Congestion Avoidance*, in "IEEE Transactions on Mobile Computing", January 2017, vol. 16, n^{o} 1, p. 58-72 [*DOI :* 10.1109/TMC.2016.2538226], https://hal.inria.fr/hal-01426424.

Articles in National Peer-Reviewed Journal

[13] N. ANCIAUX, P. PUCHERAL, M. BEHAR-TOUCHAIS, V.-L. BENABOU, G. BRUNAUX, A. LEFEVRE, N. MARTIAL-BRAZ, J. ROCHFELD, N. SAUPHANOR-BROUILLAUD, B. SCHULZ, J. SENECHAL, C. ZOLYNSKI.*Dossier Contenus Numériques Revue Contrats, Concurrence, Consommation - Contenus Numériques*, in "Contrats concurrence consommation", February 2017, https://hal.archives-ouvertes.fr/hal-01432544.

[14] C. BERTHET, C. ZOLYNSKI, N. ANCIAUX, P. PUCHERAL.*" Contenus numériques et récupération des données : un nouvel outil d' 'empouvoirement' du consommateur ? "*, in "Dalloz IP/IT", January 2017, vol. IP IT / 10, https://hal.inria.fr/hal-01429939.

[15] P. PUCHERAL, N. ANCIAUX, M. BEHAR-TOUCHAIS, V.-L. BENABOU, N. MARTIAL-BRAZ, J. ROCHFELD, N. SAUPHANOR-BROUILLAUD, B. SCHULZ, J. SENECHAL, C. ZOLYNSKI.*Dossier Contenus Numériques / Données*, in "Contrats concurrence consommation", February 2017, https://hal.inria.fr/hal-01429951.

International Conferences with Proceedings

[16] C. JACOMME, S. KREMER, G. SCERRI.*Symbolic Models for Isolated Execution Environments*, in "2nd IEEE European Symposium on Security and Privacy (EuroS&P'17)", Paris, France, C. HRIŢCU (editor), Proceedings of the 2nd IEEE European Symposium on Security and Privacy, Springer, April 2017, https://hal.inria.fr/hal-01396291.

[17] I. L. PICOLI, C. V. PASCO, B. Þ. JÓNSSON, L. BOUGANIM, P. BONNET.*uFLIP-OC: Understanding Flash I/O Patterns on Open-Channel Solid-State Drives*, in "APSys'17", Mumbai, India, September 2017, p. 1-7 [*DOI :* 10.1145/3124680.3124741], https://hal.archives-ouvertes.fr/hal-01654985.

[18] P. TRAN-VAN, N. ANCIAUX, P. PUCHERAL.*A new Sharing Paradigm for the Personal Cloud*, in "TrustBus 2017 - 14th International Conference on Trust, Privacy and Security in Digital Business", Lyon, France, August 2017, https://hal.inria.fr/hal-01675092.

[19] P. TRAN-VAN, N. ANCIAUX, P. PUCHERAL.*SWYSWYK: a new Sharing Paradigm for the Personal Cloud*, in "ADMA 2017 - International Conference on Advanced Data Mining and Applications", Singapore, Indonesia, November 2017, https://hal.inria.fr/hal-01675091.

[20] P. TRAN-VAN, N. ANCIAUX, P. PUCHERAL.*SWYSWYK: A Privacy-by-Design Paradigm for Personal Information Management Systems*, in "International Conference on Information Systems Development (ISD)", Cyprus, Cyprus, September 2017, https://hal.inria.fr/hal-01675090.

[21] P. TRAN-VAN, N. ANCIAUX, P. PUCHERAL.*Reconciling Privacy and Data Sharing in a Smart and Connected Surrounding*, in "International Conference on Extending Database Technology (EDBT)", Vienna, Austria, March 2018, https://hal.inria.fr/hal-01675093.

Conferences without Proceedings

[22] A. MICHEL, B. NGUYEN, P. PUCHERAL.*Managing Distributed Queries under Personalized Anonymity Constraints*, in "DATA", MADRID, Spain, 2017, https://hal.archives-ouvertes.fr/hal-01682316.

Scientific Books (or Scientific Book chapters)

[23] L. BOUGANIM.*Data Skew*, in "Encyclopedia of Database Systems (2nd edition)", L. LIU, M. O¨ZSU (editors), Springer, 2017 [*DOI :* 10.1007/978-1-4899-7993-3_1088-2], https://hal.archives-ouvertes.fr/hal-01656691.

[24] L. BOUGANIM.*Query Load Balancing in Parallel Database Systems*, in "Encyclopedia of Database Systems (2nd edition)", L. LIU, M. O¨ZSU (editors), Springer, 2017, p. 1-6 [*DOI :* 10.1007/978-1-4899-7993-3_1080-2], https://hal.inria.fr/hal-01660649.

References in notes

[25] S. LALLALI, N. ANCIAUX, I. SANDU POPA, P. PUCHERAL.*A Secure Search Engine for the Personal Cloud*, in "Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data (SIGMOD '15). Demo paper", Melbourne, Australia, 2015, p. 1445-1450 [*DOI :* 10.1145/2723372.2735376], https://hal.inria.fr/hal-01176473.

Project-Team POEMS

Wave propagation: mathematical analysis and simulation

IN PARTNERSHIP WITH:
CNRS

Ecole nationale supérieure des techniques avancées

RESEARCH CENTER
Saclay - Île-de-France

THEME
Numerical schemes and simulations

Table of contents

Project-Team POEMS

Creation of the Project-Team: 2005 January 01, end of the Project-Team: 2017 December 31

Keywords:

Computer Science and Digital Science:

A6. - Modeling, simulation and control
A6.1. - Mathematical Modeling
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.4. - Multiscale modeling
A6.1.5. - Multiphysics modeling
A6.1.6. - Fractal Modeling
A6.2. - Scientific Computing, Numerical Analysis & Optimization
A6.2.1. - Numerical analysis of PDE and ODE
A6.2.7. - High performance computing
A6.3.1. - Inverse problems

Other Research Topics and Application Domains:

B4.1. - Fossile energy production (oil, gas)
B5.3. - Nanotechnology
B5.4. - Microelectronics
B9.4.2. - Mathematics
B9.4.3. - Physics
B9.9.1. - Environmental risks

1. Personnel

Research Scientists

Anne-Sophie Bonnet-Ben Dhia [Team leader, CNRS, Senior Researcher, HDR]
Eliane Bécache [Inria, Researcher, HDR]
Marc Bonnet [CNRS, Researcher, HDR]
Stéphanie Chaillat [CNRS, Researcher]
Christophe Hazard [CNRS, Researcher, HDR]
Patrick Joly [Inria, Senior Researcher, HDR]
Jean-François Mercier [CNRS, Researcher, HDR]
Axel Modave [CNRS, Researcher]
Maryna Kachanovska [Inria, Researcher, from Oct 2017]

Faculty Members

Marc Lenoir [École Nationale Supérieure de Techniques Avancées, Professor, HDR]
Laurent Bourgeois [École Nationale Supérieure de Techniques Avancées, Professor, HDR]
Patrick Ciarlet [École Nationale Supérieure de Techniques Avancées, Professor, HDR]
Sonia Fliss [École Nationale Supérieure de Techniques Avancées, Professor]
Ennio Fedrizzi [École Nationale Supérieure de Techniques Avancées, Professor, from Sep 2017]
Eric Lunéville [École Nationale Supérieure de Techniques Avancées, Professor]
Konstantin Pankrashkin [Univ Paris-Sud, Associate Professor, until Aug 2017]

Technical Staff

Colin Chambeyron [CNRS, Engineer]
Nicolas Kielbasiewicz [CNRS, Engineer]
Christophe Mathulik [École Nationale Supérieure de Techniques Avancées]

PhD Students

Emile Parolin [Inria, from Sep 2017]
Zouhair Adnani [PhD Student]
Antoine Bensalah [PhD Student]
Antoine Bera [PhD Student]
Luca Desiderio [PhD Student, until Jan 2017]
Léandre Giret [PhD Student]
Sandrine Paolantoni [PhD Student]
Laure Pesudo [PhD Student, until Oct 2017]
Arnaud Recoquillay [PhD Student]
Yohanes Tjandrawidjaja [PhD Student]
Emmanuel Zerbib [PhD Student,until Feb 2017]
Clément Beneteau [PhD Student,from Oct 2017]
Hajer Methenni [PhD Student,from Oct 2017]
Yacine Abourrig [PhD Student,from Oct 2017]
Damien Mavaleix [PhD Student,from Oct 2017]

Post-Doctoral Fellows

Maryna Kachanovska [Autre entreprise publique, until Sep 2017]
Nicolas Salles [École Nationale Supérieure de Techniques Avancées, until Aug 2017]
Laure Pesudo [École Nationale Supérieure de Techniques Avancées, from Oct 2017]
Felix Kpadonou [ENSTA, from Oct 2017]
Dmitry Ponomarev [ENSTA]
Faisal Amlani [ENSTA]

Administrative Assistants

Natalia Alves [Inria]
Emmanuelle Perrot [Inria]
Corinne Chen [ENSTA, Assistant]

2. Overall Objectives

2.1. The topic of waves

The propagation of waves is one of the most common physical phenomena one can meet in nature. From the human scale (sounds, vibrations, water waves, telecommunications, radar) to the scales of the universe (electromagnetic waves, gravity waves) and of the atoms (spontaneous or stimulated emission, interferences between particles), the emission and the reception of waves are our privileged way to understand the world that surrounds us.

The study and the simulation of wave propagation phenomena constitute a very broad and active field of research in various domains of physics and engineering sciences.

The variety and the complexity of the underlying problems, their scientific and industrial interest, the existence of a common mathematical structure to these problems from different areas justify together a research project in Scientific Computing entirely devoted to this theme.

2.2. POEMS activities

The project POEMS is an UMR (Unité Mixte de Recherche) between CNRS, ENSTA ParisTech and Inria (UMR 7231). The general activity of the project is oriented toward the design, the analysis, the numerical approximation and the control of mathematical models for the description of wave propagation in mechanics, physics and engineering sciences.

Beyond the general objective of contributing to the progress of the scientific knowledge, four goals can be ascribed to the project:

- the development of expertise relative to various types of waves (acoustic, elastic, electromagnetic, gravity waves, ...), their modelling and numerical simulation,
- the treatment of complex problems whose simulation is close enough to real life situations and industrial applications,
- the development of original mathematical and numerical techniques,
- the development of computational codes, in particular in collaboration with external partners (scientists from other disciplines, industry, state companies...)

3. Research Program

3.1. General description

Our activity relies on the existence of boundary value problems established by physicists to model the propagation of waves in various situations. The basic ingredient is a partial differential equation of the hyperbolic type, whose prototype is the wave equation (or the Helmholtz equation if time-periodic solutions are considered). Nowadays, the numerical techniques for solving the basic academic problems are well mastered. However, the solution of complex wave propagation problems close to real applications still raises (essentially open) problems which constitute a real challenge for applied mathematicians. In particular, several difficulties arise when extending the results and the methods from the scalar wave equation to vectorial problems modeling wave propagation in electromagnetism or elastodynamics.

A large part of research in mathematics, when applied to wave propagation problems, is oriented towards the following goals:

- The design of new numerical methods, increasingly accurate and efficient.
- The development of artificial transparent boundary conditions for handling unbounded propagation domains.
- The treatment of more and more complex configurations (non local models, non linear models, coupled systems, periodic media).
- The study of specific phenomena such as guided waves and resonances, which raise mathematical questions of spectral theory.
- The development of approximate models via asymptotic analysis with multiple scales (thin layers, boundary layers effects, small heterogeneities, homogenization, ...).
- The development and the analysis of algorithms for inverse problems (in particular for inverse scattering problems) and imaging techniques, using data from wave phenomena.

3.2. New schemes for time-domain simulations

Problems of wave propagation naturally arise as problems of evolution and it is necessary to have efficient methods for the calculation of their solution, directly in the time domain. The development and analysis of such methods has been in the past an important part of POEMS activity. Nowadays, there exists a large variety of higher order numerical methods that allow us to solve with good accuracy and in short computational time most classical wave propagation problems. However, when on wishes to deal with real life applications, one has to tackle problems which are complex in many ways: they involve multi-physics, non standard (possibly nonlinear) constitutive laws, highly heterogeneous media with high contrasts of coefficients, complex geometries... In many cases, such problems escape to the direct application of the above mentioned methods and ad hoc dedicated methods have to be designed. Such methods are most often of hybrid nature, which includes domain decomposition methods and subgridding, mixing of integral equations and PDEs, and artificial boundary conditions. In time domain, a particularly challenging issue is the time stability, in particular concerning the coupling of algorithms. To cope with this major difficulty, a key issue (and a kind of graal for numerical analysts) is the development of energy preserving methods which is one of the specificity of the research developed at POEMS in this field.

3.3. Integral equations

Our activity in this field aims at developing accurate and fast methods for 3D problems.

On one hand, we developed a systematic approach to the analytical evaluation of singular integrals, which arise in the computation of the matrices of integral equations when two elements of the mesh are either touching each other or geometrically close.

On the other hand, POEMS is developing Fast Boundary Element Methods for 3D acoustics or elastodynamics, with applications to soil-structure interaction, seismology or seismic imaging.

Finally, a posteriori error analysis methodologies and adaptivity for boundary integral equation formulations of acoustic, electromagnetic and elastic wave propagation are investigated, continuing what was intiated during the ANR project RAFFINE.

3.4. Domain decomposition methods

This is a come back to a topic in which POEMS contributed in the 1990's. It is motivated by our collaborations with the CEA-CESTA and the CEA-LIST, for the solution of large problems in time-harmonic electromagnetism and elastodynamics.

We combine in an original manner classical ideas of Domain Decomposition Methods with the specific formulations that we use for wave problems in unbounded domains, taking benefit of the available analytical representations of the solution (integral representation, modal expansion etc...).

One ANR project (NonLocalDD) supports this research.

3.5. Wave propagation in complex media

Our objective is first to develop efficient numerical approaches for the propagation of waves in heterogeneous media, taking into account their complex microstructure.

We aim on one hand to improve homogenized modeling of periodic media, by deriving enriched boundary conditions (or transmission conditions if the periodic structure is embedded in a homogeneous matrix) which take into account the boundary layer phenomena. On the other hand, we like to develop multi-scale numerical methods when the assumption of periodicity on the spatial distribution of the heterogeneities is relaxed, or even completely lost. The general idea consists in a coupling between a macroscopic solver, based on a coarse mesh, with some microscopic representation of the field. This latter can be obtained by a numerical microscopic solver or by an analytical asymptotic expansion. This leads to two very different approaches which may be relevant for very different applications.

Extraordinary phenomena regarding the propagation of electromagnetic or acoustic waves appear in materials which have non classical properties: materials with a complex periodic microstructure that behave as materials with negative physical parameters, metals with a negative dielectric permittivity at optical frequencies, magnetized plasmas endowed with a strongly anisotropic and sign-indefinite permittivity tensor. These non classical materials raise original questions from theoretical and numerical points of view.

The objective is to study the well-posedness in this unusual context where physical parameters are sign-changing. New functional frameworks must be introduced, due, for instance, to hypersingularities of the electromagnetic field which appear at corners of metamaterials. This has of course numerical counterparts. In particular, classical Perfectly Matched Layers are unstable in these dispersive media, and new approaches must be developed.

3.6. Spectral theory and modal approaches

The study of waveguides is a longstanding and major topic of the team. Concerning the selfadjoint spectral theory for open waveguides, we turned recently to the very important case of periodic media. One objective is to design periodic structures with localized perturbations to create gaps in the spectrum, containing isolating eigenvalues.

Then, we would like to go further in proving the absence of localized modes in non uniform open waveguides. An original approach has been successfully applied to the scalar problem of a waveguides junctions or bent waveguides. The challenge now is to extend these ideas to vectorial problems (for applications to electromagnetism or elastodynamics) and to junctions of periodic waveguides.

Besides, we will continue our activity on modal methods for closed waveguides. In particular, we aim at extending the enriched modal method to take into account curvature and rough boundaries.

Finally, we are developing asymptotic models for networks of thin waveguides which arise in several applications (electric networks, simulation of lung, nanophotonics...).

3.7. Inverse problems

Building on the strong expertise of POEMS in the mathematical modeling of waves, most of our contributions aim at improving inverse scattering methodologies.

We acquired some expertise on the so called Linear Sampling Method, from both the theoretical and the practical points of view. Besides, we are working on topological derivative methods, which exploit small-defect asymptotics of misfit functionals and can thus be viewed as an alternative sampling approach, which can take benefit of our expertise on asymptotic methods.

An originality of our activity is to consider inverse scattering in waveguides (the inverse scattering community generally considers only free-space configurations). This is motivated at the same time by specific issues concerning the ill-posedness of the identification process and by applications to non-destructive techniques, for waveguide configurations (cables, pipes, plates etc...).

Lastly, we continue our work on the so-called exterior approach for solving inverse obstacle problems, which associates quasi-reversibility and level set methods. The objective is now to extend it to evolution problems.

4. Application Domains

4.1. Acoustics

Two particular subjects have retained our attention recently.

1- Aeroacoustics, or more precisely, acoustic propagation in a moving compressible fluid, has been for our team a very challenging topic, which gave rise to a lot of open questions, from the modeling until the numerical approximation of existing models. Our works in this area are partially supported by EADS and Airbus. The final objective is to reduce the noise radiated by Airbus planes.

2- Musical acoustics constitute a particularly attractive application. We are concerned by the simulation of musical instruments whose objectives are both a better understanding of the behavior of existing instruments and an aid for the manufacturing of new instruments. We have successively considered the timpani, the guitar and the piano. This activity is continuing in the framework of the European Project BATWOMAN.

4.2. Electromagnetism

Applied mathematics for electromagnetism during the last ten years have mainly concerned stealth technology and electromagnetic compatibility. These areas are still motivating research in computational sciences (large scale computation) and mathematical modeling (derivation of simplified models for multiscale problems). These topics are developed in collaboration with CEA, DGA and ONERA.

Electromagnetic propagation in non classical media opens a wide and unexplored field of research in applied mathematics. This is the case of wave propagation in photonic crystals, metamaterials or magnetized plasmas.

Finally, the simulation electromagnetic (possibly complex, even fractal) networks is motivated by destructive testing applications. This topic is developed in partnership with CEA-LIST.

4.3. Elastodynamics

Wave propagation in solids is with no doubt, among the three fundamental domains that are acoustics, electromagnetism and elastodynamics, the one that poses the most significant difficulties from mathematical and numerical points of view. A major application topic has emerged during the past years : the non destructive testing by ultra-sounds which is the main topic of our collaboration with CEA-LIST. On the other hand, we are developing efficient integral equation modelling for geophysical applications (soil-structure interaction for civil engineering, seismology).

5. Highlights of the Year

5.1. Waves diffracted by Patrick Joly

On the occasion of Patrick Joly's 60th birthday, a conference with about hundred attendees has been organized by Sonia Fliss, Xavier Claeys, Bérangère Delourme and Julien Diaz, from August 28th to August 30th 2017, to acknowledge and celebrate his decisive scientific contributions in the mathematical and numerical analysis of wave propagation.

Below is the list of invited Speakers

- Grégoire Allaire (CMAP, Ecole Polytechnique)
- Jean-David Benamou (Inria Paris)
- Anne-Sophie Bonnet-BenDhia (ENSTA/CNRS/Inria POems)
- Yann Brenier (Centre de Mathématiques Laurent Schwarz)
- Antoine Chaigne (MDW Vienna, Autriche)
- Simon Chandler-Wilde (Univ. Reading, UK)
- Lucas Chesnel (Inria Defi / CMAP Ecole Polytechnique)
- Bernardo Cockburn (Univ. Minnesota, USA)
- Francis Collino (freelance)
- Alexander Comech (Vienna University, Autriche)
- Martin Costabel (IRMAR, Univ. Rennes)
- Bruno Despres (LJLL UPMC)
- Bjorn Engquist (Univ. Texas Austin, USA)

- Martin Gander (Univ. Genève Suisse)
- Marcus Grote (Univ. Bâle Suisse)
- Houssem Haddar (Inria Defi / CMAP Ecole Polytechnique)
- Laurence Halpern (LAGA Univ. Paris 13)
- Thomas Hagstrom (Southern Methodist University Dallas, USA)
- Jan Hesthaven (EPF Lauzanne Suisse)
- Ralf Hiptmair (ETH Zurich Suisse)
- Andreas Kirsch (Karlsruhe Institute of Technology Allemagne)
- Claude Le Bris (CERMICS ENPC)
- Jérome Le Rousseau (LAGA Univ. Paris 13)
- Pierre Louis Lions (College de France)
- Peter Monk (Univ. Delaware)
- Serge Nicaise (Univ. Valenciennes)
- Konstantin Pankrashkin (Univ. Paris 11 Orsay)
- George Papanicolaou (Stanford University USA)
- Jerónimo Rodriguez (Univ. Saint Jacques de Compostelle)
- Chrysoula Tsogka (University of Crete Grèce)
- Ricardo Weder (University of Mexico Mexique)

A short presentation of former PhD students of Patrick Joly has also given an overview of his recent activities:

- Antoine Bensalah (ENSTA/CNRS/Inria Poems)
- Maxence Cassier (University of Utah)
- Juliette Chabassier (Inria Bordeaux, EPI Magique 3D)
- Julien Coatleven (IFP)
- Sebastien Imperiale (Inria Saclay, EPI M3DISIM)
- Elizaveta Vasilevskaya (High School teacher)

5.2. A day for Marc Lenoir

A day entitled *Un Lenoir... ça Marc... donc ça se fête* was organized at ENSTA on June, 23th, and gathered about 60 people. This day was intended to make a festive tribute to Marc Lenoir for his role in what has become the Applied Mathematics Laboratory of ENSTA (including POEMS). Two scientific talks have been given by longtime friends of Marc: Michel Crouzeix (University of Rennes) and Jacques Rappaz (Ecole Polytechnique Fédérale de Lausanne, Switzerland). The other talks, which emphasized the scientific and human qualities of Marc, were given by four former students: Nicolas Salles, Eric Lunéville and Christophe Hazard (all from POEMS) and Nabil Gmati (LAMSIN, Tunis).

6. New Software and Platforms

6.1. COFFEE

FUNCTIONAL DESCRIPTION: COFFEE is a 3D BEM-accelerated FMM solver for linear elastodynamics (full implementation, 30 000 lines of Fortran 90). The 3-D elastodynamic equations are solved with the boundary element method accelerated by the multi-level fast multipole method. The fundamental solutions for the infinite space are used in this implementation. A boundary element-boundary element coupling strategy is also implemented so multi-region problems (strata inside a valley for example) can be solved.

- Contact: Stéphanie Chaillat

6.2. XLiFE++

FUNCTIONAL DESCRIPTION: XLiFE++ is a Finite Element library in C++ based on philosophy of the previous library Melina in Fortran but with new capabilities (boundary elements, discontinuous Galerkin methods, more integrated tools -in particular mesh tools - and high performance computing skills, multithread and GPU computation).

- Contact: Eric Lunéville

7. New Results

7.1. New schemes for time-domain simulations

7.1.1. Solving the Homogeneous Isotropic Linear Elastodynamics Equations Using Potentials

Participant: Patrick Joly.

This work is done in collaboration with Sébastien Impériale (EPI M3DISIM) and Jorge Albella from the University of Santiago de Compostela. We consider the numerical solution of 2D elastodynamic equations using the decomposition of the displacement fields into potentials. This appears as a challenge for finite element methods. We address here the particular question of free boundary conditions. A stable (mixed) variational formulation of the evolution problem is proposed based on a clever choice of Lagrange multipliers. This is expected to be efficient when the velocity of shear waves is much smaller than the velocity of pressure waves, since one can adapt the discretization to each type of waves.

7.1.2. Discontinuous Galerkin method with high-order absorbing boundary conditions

Participant: Axel Modave.

This work is done in collaboration with Andreas Atle from TOTAL, Jesse Chan from Rice University and Tim Warburton from Virginia Tech.

Discontinuous Galerkin finite element schemes exhibit attractive features for large-scale time-domain wave-propagation simulations on modern parallel architectures (e.g. GPU clusters). For many applications, these schemes must be coupled with non-reflective boundary treatments to limit the size of the computational domain without losing accuracy or computational efficiency, which remains a challenging task.

We propose a combination of a nodal discontinuous Galerkin method with high-order absorbing boundary conditions (HABCs) for cuboidal computational domains. Compatibility conditions are derived for HABCs intersecting at the edges and the corners of a cuboidal domain. We propose a GPU implementation of the computational procedure, which results in a multidimensional solver with equations to be solved on 0D, 1D, 2D and 3D spatial regions. Numerical results demonstrate both the accuracy and the computational efficiency of our approach.

7.2. Integral equations

7.2.1. Mesh adaptation for the fast multipole method in acoustics

Participants: Faisal Amlani, Stéphanie Chaillat.

This work is done in collaboration with Adrien Loseille (EPI Gamma3). We introduce a metric-based anisotropic mesh adaptation strategy for the fast multipole accelerated boundary element method (FM-BEM) applied to exterior boundary value problems of the three-dimensional Helmholtz equation. The present methodology is independent of discretization technique and iteratively constructs meshes refined in size, shape and orientation according to an *optimal* metric reliant on a reconstructed Hessian of the boundary solution. The resulting adaptation is anisotropic in nature and numerical examples demonstrate optimal convergence rates for domains that include geometric singularities such as corners and ridges.

7.2.2. Coupling integral equations and high-frequency methods

Participants: Marc Bonnet, Marc Lenoir, Eric Lunéville, Laure Pesudo, Nicolas Salles.

This theme concerns wave propagation phenomena which involve two different space scales, namely, on the one hand, a medium scale associated with lengths of the same order of magnitude as the wavelength (medium-frequency regime) and on the other hand, a long scale related to lengths which are large compared to the wavelength (high-frequency regime). Integral equation methods are known to be well suited for the former, whereas high-frequency methods such as geometric optics are generally used for the latter. Because of the presence of both scales, both kinds of simulation methods are simultaneously needed but these techniques do not lend themselves easily to coupling.

A first situation, considered by Marc Lenoir, Eric Lunéville and Nicolas Salles, is the scattering of an acoustic wave by two sound-hard obstacles: a large obstacle subject to high-frequency regime relatively to the wavelength and a small one subject to medium-frequency regime. The technique proposed in this case consists in an iterative method which allows to decouple the two obstacles and to use Geometric Optics for the large obstacle and Boundary Element Method for the small obstacle. The method is implemented on the XLife++ library developed in the lab.

The second situation, undertaken in the context of the PhD thesis of Laure Pesudo, is the subject of a partnership with CEA LIST and a collaboration with Francis Collino. Modelling ultrasonic non destructive testing (NDT) experiments simultaneously involves the scattering of waves by defects of moderate size (for which discretization-based methods such as the BEM are appropriate) and propagation over large distances (requiring high-frequency approximations). A new hybrid strategy between the boundary element method (BEM) and ray tracing is proposed in order to allow the accurate and quick simulation of high frequency Non Destructive Testing (NDT) configurations involving diffraction phenomena. Results from its implementation to 2D acoustic NDT-like diffraction configurations have been obtained. The strategy proposed is however generic, and can be extended to three-dimensional configurations and elastodynamic wave propagation.

7.2.3. Dynamic soil-structure interaction

Participants: Marc Bonnet, Stéphanie Chaillat, Zouhair Adnani.

This work, undertaken in the context of the PhD thesis of Zouhair Adnani (CIFRE partnership with EDF), concerns the simulation of dynamic soil-structure interaction (SSI) in connection with seismic assessment of civil engineering structures. Because of the complementary specificities of the finite element method (FEM) and the boundary element method (BEM), it is natural to use the BEM to model the unbounded soil domain, while the FEM is applied for the bounded region comprising the structure undergoing assessment, and possibly its close-range soil environment.

The originality of this work is to formulate, implement, and evaluate on realistic test examples, a computational strategy that combines the fast multipole accelerated boundary element method (visco-elastodynamic COFFEE solver), and the EDF in-house FEM code Code_Aster. In a preliminary phase, the evaluation of transient elastodynamic responses via the Fourier synthesis of frequency domain solutions computed using COFFEE (see Section 5.1) has been studied on several test problems, achieving substantial improvements of computational efficiency for this component of SSI analysis.

The coupling between the two methods is then done in a black-box fashion with the substructuring method by computing the soil impedance (i.e. elastodynamic Poincaré-Steklov) operator relating forces to displacements on the FEM-BEM coupling interface. One of the main challenges is that this operator cannot be assembled due to the iterative nature of the FM-BEM and the potentially large number of degrees of freedom supported by the interface. To reduce the computational costs, we instead compute its projection on a reduced basis of interface modes, which requires to perform as many FM-BEM calculations as interface modes selected. This approach has so far been compared to reference solutions and validated for superficial and buried foundations on homogeneous or heterogeneous soil.

7.2.4. Volume Integral Formulations

Participant: Marc Bonnet.

Volume integral equations (VIEs), also known as Lippmann-Schwinger integral equations, arise naturally when considering the scattering of waves by penetrable, and possibly heterogeneous, inhomogeneities embedded in a homogeneous background medium (for which a fundamental solution is explicitly known). Their derivation and use in e.g. acoustics, elastodynamics or electromagnetism goes back several decades. Since their geometrical support is confined to the spatial region where material properties differ from the background, VIEs are in particular useful for the derivation and justification of homogenized or asymptotic models (the latter providing our main motivation for this study, in connection with [section gradient topologique]). By directly linking remote measurements to unknown inhomogeneities, VIEs also provide a convenient forward modeling approach for medium imaging inverse problems. However, whereas the theory of boundary integral equations is extensively documented, the mathematical properties of VIEs have undergone a comparatively modest coverage, much of it pertaining to electromagnetic scattering problems.

In this work, we investigate the solvability of VIE formulations arising in elastodynamic scattering by penetrable obstacles. The elasticity tensor and mass density are allowed to be smoothly heterogeneous inside the obstacle and may be discontinuous across the background-obstacle interface, the background elastic material being homogeneous. Both materials may be anisotropic, within certain limitations for the background medium.

Towards this goal, we have introduced a modified version of the singular volume integral equation (SVIE) governing the corresponding elastostatic (i.e. zero frequency) problem, and shown it to be of second kind involving a contraction operator, i.e. solvable by Neumann series, for any background material and inhomogeneity material and geometry. Then, the solvability of VIEs for frequency-domain elastodynamic scattering problems follows by a compact perturbation argument, assuming uniqueness to be established. In particular, in an earlier work, we have established a uniqueness result for the anisotropic background case (where, to avoid difficulties associated with existing radiation conditions for anisotropic elastic media, we have proposed an alternative definition of the radiating character of solutions, which is equivalent to the classical Sommerfeld-Kupradze conditions for the isotropic background case). This investigation extends work by Potthast (1999) on 2D electromagnetic problems (transverse-electric polarization conditions) involving orthotropic inhomogeneities in a isotropic background, and contains recent results on the solvability of Eshelby's equivalent inclusion problem as special cases. The proposed modified SVIE is also useful for fixed-point iterative solution methods, as Neumannn series then converge (i) unconditionally for static problems and (ii) on some inhomogeneity configurations for which divergence occurs with the usual SVIE for wave scattering problems.

7.3. Domain decomposition methods

7.3.1. Transparent boundary conditions with overlap in unbounded anisotropic media

Participants: Anne-Sophie Bonnet-Ben Dhia, Sonia Fliss, Yohanes Tjandrawidjaja.

This work is done in the framework of the PhD of Yohanes Tjandrawidjaja, funded by CEA-LIST, in collaboration with Vahan Baronian form CEA. This follows the PhD of Antoine Tonnoir (now Assistant Professor at Insa of Rouen) who developed a new approach, the Half-Space Matching Method, to solve scattering problems in 2D unbounded anisotropic media. The objective is to extend the method to a 3D plate of finite width.

In 2D, our approach consists in coupling several plane-waves representations of the solution in half-spaces surrounding the defect with a FE computation of the solution around the defect. The difficulty is to ensure that all these representations match, in particular in the infinite intersections of the half-spaces. It leads to a Fredholm formulation which couples, via integral operators, the solution in a bounded domain including the defect and some traces of the solution on the edge of the half-planes.

The extension to 3D elastic plates requires some generalizations of the formulation which are not obvious. In particular, we have to use Neumann traces of the solution, which raises difficult theoretical questions.

As a first step, we have considered a scattering problem outside a convex polygonal scatterer for a general class of boundary conditions, using the Half-Space Matching Method. Using the Mellin Transform, we are able to show that this system is coercive + compact in presence of dissipation. We have also proved the convergence of the discrete method with respect to the size of truncation of the Fourier integrals, and with respect to the mesh size. This is the object of a paper that has been submitted.

In parallel, the main ingredient for the numerical method in 3D has been developed. It is the modal/Fourier representation of the elastic field in a semi-infinite plate, as a function of the trace of the displacement and of the normal stress. This has been done in the isotropic case.

7.3.2. Domain Decomposition Methods for the neutron diffusion equation

Participants: Patrick Ciarlet, Léandre Giret.

This work is done in collaboration with Erell Jamelot (CEA-DEN, Saclay) and Félix Kpadonou (LMV, UVSQ). Studying numerically the steady state of a nuclear core reactor is expensive, in terms of memory storage and computational time. In its simplest form, one must solve a neutron diffusion equation with low-regularity solutions, discretized by mixed finite element techniques, within a loop. Iterating in this loop allows to compute the smallest eigenvalue of the system, which determines the critical, or non-critical, state of the core. This problem fits within the framework of high performance computing so, in order both to optimize the memory storage and to reduce the computational time, one can use a domain decomposition method, which is then implemented on a parallel computer: this is the strategy used for the APOLLO3 neutronics code. The development of non-conforming DD methods for the neutron diffusion equation with low-regularity solutions has recently been finalized, cf. [PC,EJ,FK'1x]. The theory for the eigenvalue problem is also understood. The current research now focuses on the numerical analysis of the full suite of algorithms to prove convergence for the complete multigroup SPN model (which involves coupled diffusion equations).

7.4. Wave propagation in complex media

7.4.1. Perfectly Matched Layers in plasmas and metamaterials

Participants: Eliane Bécache, Maryna Kachanovska.

In this work we consider the problem of the modelling of 2D anisotropic dispersive wave propagation in unbounded domains with the help of perfectly matched layers (PML). We study the Maxwell equations in passive media with the frequency-dependent diagonal tensor of dielectric permittivity and magnetic permeability. An application of the traditional PMLs to this kind of problems often results in instabilities, due to the presence of so-called backward propagating waves. In previous works, this instability was overcome with the help of the frequency-dependent correction of the PML, for isotropic dispersive models.

We show that this idea can be extended to a more general class of models (uniaxial cold plasma, some anisotropic metamaterials). Crucially, we base our considerations on the Laplace-domain techniques. This allows to avoid the analysis of the group and phase velocity (used before) but study (rather formally) coercivity properties of the sesquilinear form corresponding to the PML model in the Laplace domain. The advantage of this method is that it permits to treat problems with dissipation, and provides an intuition on how to obtain explicit energy estimates for the resulting PML models in the time domain. However, such analysis does not allow to obtain easily the necessary stability condition of the PML. We demonstrate that the necessary stability conditions of the PML can be rewritten for a class of models in a form that is easy to verify, and demonstrate that these conditions are sufficient for the stability of the new PMLs with the help of the Laplace-domain techniques. Thanks to the Laplace domain analysis, we are able to rewrite a PML system in the time domain in a form, for which the derivation of energy estimates is simplified (compared to other formulations).

7.4.2. Transparent Boundary Conditions for the Wave Propagation in Fractal Trees

Participants: Patrick Joly, Maryna Kachanovska.

This work, done in collaboration with Adrien Semin (Postdoctoral student at the Technische Universität of Berlin), is dedicated to an efficient resolution of the wave equation in self-similar trees (e.g. wave propagation in a human lung). In this case it is possible to avoid computing the solution at deeper levels of the tree by using the transparent boundary conditions. The corresponding DtN operator is defined by a functional equation in the frequency domain. In this work we propose and compare two approaches to the discretization of this operator in the time domain. The first one is based on the multistep convolution quadrature, while the second one stems from the rational approximations.

7.4.3. High order transmission conditions between homogeneous and homogenized periodic half-spaces

Participants: Sonia Fliss, Clément Beneteau.

This work is a part of the PhD of Valentin Vinoles, and is done in collaboration with Xavier Claeys from Paris 6 University and EPI ALPINE. It is motivated by the fact that classical homogenization theory poorly takes into account interfaces, which is particularly unfortunate when considering negative materials, because important phenomena arise precisely at their surface (plasmonic waves for instance). To overcome this limitation, we want to construct high order transmission conditions. For now, we have treated the case of a plane interface between a homogeneous and a periodic half spaces. Using matched asymptotic techniques, we have derived high order transmission conditions. We have then introduced an approximate model associated to this asymptotic expansions which consists in replacing the periodic media by an effective one but the transmission conditions are not classical. The obtained conditions involve Laplace- Beltrami operators at the interface and requires to solve cell problems in periodicity cell (as in classical homogenisation) and in infinite strips (to take into account the phenomena near the interface). We establish well posedness for the approximate and error estimate which justify that this new model is more accurate near the interface and in the bulk. From a numerical point of view, the only difficulty comes from the problems set in infinite strips (one half is homogeneous and the other is periodic). This is overcome using DtN operators corresponding to the homogeneous and the periodic media. The numerical results confirm the theoretical ones.

7.5. Spectral theory and modal approaches for waveguides

7.5.1. Modal analysis of electromagnetic dispersive media

Participants: Christophe Hazard, Sandrine Paolantoni.

We investigate the spectral effects of an interface between a usual dielectric and a negative-index material (NIM), that is, a dispersive material whose electric permittivity and magnetic permeability become negative in some frequency range. We consider here an elementary situation, namely, 1) the simplest existing model of NIM : the Drude model (for which negativity occurs at low frequencies); 2) a two-dimensional scalar model derived from the complete Maxwell's equations; 3) the case of a simple bounded cavity: a camembert-like domain partially

lled with a portion of non dissipative Drude material. Because of the frequency dispersion (the permittivity and permeability depend on the frequency), the spectral analysis of such a cavity is unusual since it yields a nonlinear eigenvalue problem. Thanks to the use of an additional unknown, we show how to linearize the problem and we present a complete description of the spectrum.

7.5.2. Formulation of invisibility in waveguides as an eigenvalue problem

Participants: Antoine Bera, Anne-Sophie Bonnet-Ben Dhia.

This work is done in collaboration with Lucas Chesnel from EPI DEFI and Vincent Pagneux from Laboratoire d'Acoustique de l'Université du Maine. A scatterer placed in an infinite waveguide may be *invisible* at particular discrete frequencies. We consider two different definitions of invisibility: no reflection (but possible conversion or phase shift in transmission) or perfect invisibility (the scattered field is exponentially decaying at infinity). Our objective is to show that the invisibility frequencies can be characterized as eigenvalues of some spectral problems. Two different approaches are used for the two different definitions of invisibility, leading to non-selfadjoint eigenvalue problems.

More precisely, for the case of no-reflection, we define a new complex spectrum which contains as real eigenvalues the frequencies where perfect transmission occurs and the frequencies corresponding to trapped modes. In addition, we also obtain complex eigenfrequencies which can be exploited to predict frequency ranges of good transmission. Our approach relies on a simple but powerful idea, which consists in using PMLs in an original manner: while in usual PMLs the same stretching parameter is used in the inlet and the outlet, here we take them as two complex conjugated parameters. As a result, they select ingoing waves in the inlet and outgoing waves in the outlet, which is exactly what arises when the transmission is perfect. This simple idea works very well, and provides useful information on the transmission qualities of the system, much faster than any traditional approach.

7.5.3. Transparent boundary conditions for general waveguide problems

Participants: Anne-Sophie Bonnet-Ben Dhia, Sonia Fliss.

In this work, done in collaboration with Antoine Tonnoir from INSA of Rouen, we propose a construction of transparent boundary conditions which can be used for quite general waveguide problems. Classical Dirichlet-to-Neumann maps used for homogeneous acoustic waveguides can be constructed using separation of variables and the orthogonality of the modes on one transverse section. These properties are also important for the mathematical and numerical analysis of problems involving DtN maps. However this framework does not extend directly to stratified, anisotropic or periodic waveguides and for Maxwell's or elastic equations. The difficulties are that (1) the separation of variables is not always possible and (2) the modes of the waveguides are not necessarily orthogonal on the transverse section. We propose an alternative to the DtN maps which uses two artificial boundaries and is constructed using a more general orthogonality property.

7.6. Inverse problems

7.6.1. Linear Sampling Method with realistic data in waveguides

Participants: Laurent Bourgeois, Arnaud Recoquillay.

Our activities in the field of inverse scattering in waveguides with the help of sampling methods has now a quite long history. We now intend to apply these methods in the case of realistic data, that is surface data in the time domain. This is the subject of the PhD of Arnaud Recoquillay. It is motivated by Non Destructive Testing activities for tubular structures and is the object of a partnership with CEA List (Vahan Baronian).

Our strategy consists in transforming the time domain problem into a multi-frequency problem by the Fourier transform. This allows us to take full advantage of the established efficiency of modal frequency-domain sampling methods. We have already proved the feasibility of our approach in the 2D acoustic and 2D elastic case. In particular, we have shown how to optimize the number of sources/receivers and the distance between them in order to obtain the best possible identification result. Experiments are currently carried in CEA.

7.6.2. The "exterior approach" to solve inverse obstacle problems

Participants: Laurent Bourgeois, Arnaud Recoquillay.

We consider some inverse obstacle problems in acoustics by using a single incident wave, either in the frequency or in the time domain. When so few data are available, a Linear Sampling type method cannot be applied. In order to solve those kinds of problem, we propose an "exterior approach", coupling a mixed formulation of quasi-reversibility and a simple level set method. In such iterative approach, for a given defect D, we update the solution u with the help of a mixed formulation of quasi-reversibility while for a given solution u, we update the defect D with the help of a level set method based on a Poisson problem. The case of data in the frequency domain has been studied for the waveguide geometry. We currently investigate the case of data in a finite time domain.

7.6.3. A continuation method for building large invisible obstacles in waveguides

Participants: Antoine Bera, Anne-Sophie Bonnet-Ben Dhia.

In collaboration with Lucas Chesnel (EPI DEFI) and Sergei Nazarov (Saint-Petersburg University), we consider time harmonic acoustic problems in waveguides. We are interested in finding localized perturbations of a straight waveguide which are not detectable in the far field, as they produce neither reflection nor conversion of propagative modes. In other words, such *invisible* perturbation produces a scattered field which is exponentially decaying at infinity in the two infinite outlets of the waveguide.

In our previous contributions, we found a way to build smooth and small perturbations of the boundary which were almost invisible, in the sense that they were producing (in the monomode regime) no reflexions but maybe a phase shift in transmission.

The method is constructive and has been validated numerically. But the drawback is that it is limited to low frequency and small perturbations. During the last year, we have shown that the previous idea can be combined with a continuation method, in order to get larger invisible perturbations at higher frequency.

7.7. Aeroacoustics

7.7.1. Time-harmonic acoustic scattering in a vortical flow

Participants: Antoine Bensalah, Patrick Joly, Jean-François Mercier.

This activity is done in the framework of the PhD of Antoine Bensalah, in partnership with Airbus Group. We study the time-harmonic acoustic radiation in a fluid in a general flow which is not curl free, but has restricted vortical areas. The objective is to take into account the complicated coupling between acoustics and hydrodynamics. The Galbrun approach developed previously in 2D is too expensive in terms of degrees of freedom for 3D simulations. As an alternative, we propose to consider instead the Goldstein equations, which are vectorial only in the vortical areas and remain scalar elsewhere.

To begin with, we aim at determining the acoustic field radiated in 2D by a time-harmonic source in a fluid in flow. Goldstein's equations are proved to be well-posed outside a spectrum of frequencies corresponding to resonant streamlines. This band spectrum is explicitly determined for two simple geometries (an annular domain and a rectangular one with periodic conditions). Then the full model is shown to be well-posed under a coercivity condition, implying a subsonic flow with a small enough vorticity.

7.7.2. Propagation of solitons through Helmholtz resonators

Participant: Jean-François Mercier.

With Bruno Lombard (Laboratoire de Mécanique et Acoustique of Marseille), we study the propagation of nonlinear solitary acoustic waves in a 1D waveguide connected to a lattice of Helmholtz resonators. We start from an homogenized model of the literature, consisting of two coupled equations evolution: a nonlinear PDE describing acoustic waves (similar to the Burgers equation), and a linear ODE describing oscillations in the Helmholtz resonators. We have already developed a numerical modeling of this model and we have compared simulations with experimental data.

The drawback of the homogenized model is that all the resonators must be the same. In particular the reflection of an incident wave by a defect cannot be considered. To remedy this limitation, we have proposed an extension of the model, predicting two-way propagation across variable resonators. Thanks to a new discrete description of the resonators, the improved model takes into account two important features: resonators of different strengths and back-scattering effects. Comparisons with experimental data show that a closer agreement is obtained.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Contract POEMS-DGA
Participants: Eric Lunéville, Marc Lenoir, Séphanie Chaillat, Nicolas Kielbasiewicz, Nicolas Salles.

Start : 2015, End : 2018. Administrator : ENSTA.
This contract is in partnership with François Alouges and Matthieu Aussal (CMAP, Ecole Polytechnique) and concerns the improvement of Boundary Element Methods for wave propagation problems.

Contract POEMS-CEA-LIST
Participants: Marc Bonnet, Laure Pesudo.

Start : 12/01/2014, End : 11/31/2017. Administrator : CNRS.
This contract is about the coupling between high frequency methods and integral equation

Contract POEMS-EDF
Participants: Stéphanie Chaillat, Marc Bonnet, Zouhair Adnani.

Start : 12/01/2014, End : 11/31/2017. Administrator : CNRS.
This contract is about fast solvers to simulate soil-structure interactions.

9. Partnerships and Cooperations

9.1. Regional Initiatives

The post-doc of Maryna Kachanovska was funded by the Fondation Mathématique Jacques Hadamard (FMJH).

9.2. National Initiatives

9.2.1. ANR

- ANR project *RAFFINE: Robustesse, Automatisation et Fiabilité des Formulations INtégrales en propagation d'ondes : Estimateurs a posteriori et adaptivité*
  Partners: EADS, IMACS, ONERA, Thales
  Start : January 2013. End : June 2017. Administrator : Inria. Coordinator: Marc Bonnet.
- ANR project *Non-Local Domain Decomposition Methods in Electromagnetism.*
  Partners: Inria Alpines, Inria POEMS, Inria Magique 3D.
  Start : 2015, End : 2019. Administrator : Inria. Coordinator: Xavier Claeys.

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

9.3.1.1. BATWOMAN

Type: FP7 Marie Curie

Objectif: Basic Acoustics Training - & Workprogram On Methodologies for Acoustics - Network

Duration: September 2013 - August 2017

Coordinator: Martin Wifling, VIRTUAL VEHICLE (AT)

Inria contact: P. Joly

Abstract: The BATWOMAN ITN aims at structuring research training in basic and advanced acoustics and setting up a work program on methodologies for acoustics for skills development in a highly diverse research field offering multiple career options.

9.4. International Initiatives

9.4.1. Inria International Partners

9.4.1.1. Informal International Partners

Wilkins Aquino (Duke University)

Juan Pablo Borthagaray (Univ. of Maryland, College Park, USA)

Fioralba Cakoni (University of Rutgers)

Maxence Cassier (Columbia University)

Camille Carvalho (UC Merced, Merced, USA)

Christophe Geuzaine (Université de Liège)

Bojan Guzina (University of Minnesota)

Marcus Grote (Universitaet Basel)

Sergei Nazarov (Saint-Petersburg University)

Jeronimo Rodriguez (University of Santiago de Compostela)

Adrien Semin (BTU Cottbus)

Ricardo Weder (Universidad Nacional Autonoma, Mexico)

Shravan Veerapaneni (Univ. of Michigan at Ann Arbor, USA)

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Advisory and management activities

- P. Joly is a member of the scientific committee of CEA-DAM.
- E. Lunéville is the Head of UMA (Unité de Mathématiques Appliquées) at ENSTA ParisTech.

10.1.2. Scientific events organisation end selection

- E. Bécache, A. S. Bonnet-Ben Dhia, M. Bonnet, S. Fliss, C. Hazard, P. Joly and E. Lunéville were members of the scientific committee for the 13rd international conference on mathematical and numerical aspects of wave propagation (WAVES 2017), which took place in Minneapolis in May 2017.
- A. Modave co-organized with B. Thierry (CNRS/UPMC) a Young Researchers' Meeting on "Solving Large-Scale Time-Harmonic Wave Problems", which took place at UPMC on November, 2017. There were 8 talks, 2 tutorials and 20 participants, including young researchers from the Inria teams Magique3D, HiePACS and Alpines. The meeting was funded in part by the SMAI through a BOUM project.

10.1.3. Journal

- A. S. Bonnet-Ben Dhia is associate editor of SINUM (SIAM Journal of Numerical Analysis) and SIAP (SIAM Journal of Applied Mathematics).
- M. Bonnet is associate editor of Engineering Analysis with Boundary Elements
- M. Bonnet is in the editorial board of Inverse Problems.

- M. Bonnet is in the editorial board of Computational Mechanics.
- M. Bonnet is in the editorial board of Journal of Optimization Theory and Application.
- P. Ciarlet is an editor of CAMWA (Computers & Mathematics with Applications).
- P. Ciarlet is an editor of ESAIM:M2AN (Mathematical Modeling and Numerical Analysis).
- P. Joly is an editor of ESAIM:M2AN (Mathematical Modeling and Numerical Analysis).
- P. Joly is a member of the editorial board of AAMM (Advances in Applied Mathematics and Mechanics).
- P. Joly is a member of the Book Series Scientific Computing of Springer Verlag.
- The team members regularly review papers for many international journals.

10.2. Teaching - Supervision

10.2.1. Teaching

Eliane Bécache

- *Méthode des éléments finis*, ENSTA ParisTech (2nd year)
- *Compléments sur la méthode des éléments finis*, ENSTA ParisTech, (2nd year)
- *Fonctions d'une variable complexe*, ENSTA ParisTech (1st year)
- *Résolution des problèmes de diffraction par équations intégrales*, ENSTA ParisTech (3rd year) and Master "Modeling and Simulation" (M2)

Marc Bonnet

- *Problèmes inverses*, Master MS2SC (Centrale Paris and ENS Cachan)
- *Méthodes intégrales*, Master TACS (ENS Cachan)
- *Outils élémentaires d'analyse pour les équations aux dérivées partielles*, ENSTA ParisTech (1st year)

Anne-Sophie Bonnet-Ben Dhia

- *Fonctions d'une variable complexe*, ENSTA ParisTech (1st year)
- *Propagation dans les guides d'ondes*, ENSTA ParisTech (3rd year) and Master "Modeling and Simulation" (M2)
- *Non Destructive Testing*, Master "Acoustics" (M2)
- *Théorie spectrale des opérateurs autoadjoints et applications aux guides optiques*, ENSTA ParisTech (2nd year)

Laurent Bourgeois

- *Outils élémentaires pour l'analyse des équations aux dérivées partielles*, ENSTA ParisTech (1st year)
- *Fonction d'une variable complexe*, ENSTA ParisTech (1st year)

Stéphanie Chaillat

- *Introduction à la discrétisation des équations aux dérivées partielles*, ENSTA ParisTech (1st year)
- *Fonctions d'une variable complexe*, ENSTA ParisTech (1st year)
- *Equations intégrales et multipôles rapides*, Ecole doctorale MODES (Univ. Paris Est, Marne la Vallée)
- *Résolution des problèmes de diffraction par équations intégrales*, ENSTA ParisTech (3rd year) and Master "Modeling and Simulation" (M2)

Colin Chambeyron

- *Analyse réelle: optimisation libre et sous contraintes*, Dauphine University (1st year)
- *Outils mathématiques*, Dauphine University (1st year)
- *Algèbre linéaire*, Dauphine University (2nd year)

Patrick Ciarlet

- Advanced Finite Element Methods, ENSTA ParisTech (2nd year)
- Parallel Scientific Computing, ENSTA ParisTech (3rd year), and Master "Analysis, Modelling, Simulation" (M2)
- Mathematical Models and their Discretisation in Electromagnetism, ENSTA ParisTech (3rd year), and Master "Analysis, Modelling, Simulation" (M2)
- Deputy Head of the Master's Program Analysis, Modelling, Simulation, Paris-Saclay University

Sonia Fliss

- *Méthode des éléments finis*, ENSTA ParisTech (2nd year)
- *Introduction à la discrétisation des équations aux dérivées partielles*, ENSTA ParisTech (1st year).
- *Propagation des ondes dans les milieux périodiques*, ENSTA ParisTech (3rd year) and Master "Analysing, Modeling and Simulation" (M2)
- *Homogénéisation périodique*, Masters ANEDP, M4S et AMS "Analysing, Modeling and Simulation" (M2)

Christophe Hazard

- *Outils élémentaires d'analyse pour les équations aux dérivées partielles*, ENSTA ParisTech (1st year)
- *Théorie spectrale des opérateurs autoadjoints et applications aux guides optiques*, ENSTA ParisTech (2nd year)

Patrick Joly

- *Introduction à la discrétisation des équations aux dérivées partielles*, ENSTA ParisTech (1st year)
- *Propagation des ondes dans les milieux périodiques*, ENSTA ParisTech (3rd year) and Master "Modeling and Simulation" (M2)

Nicolas Kielbasiewicz

- *Programmation scientifique et simulation numérique*, ENSTA ParisTech (2nd year)
- *Parallélisme et calcul réparti*, ENSTA ParisTech (Master 2)

Marc Lenoir

- *Fonctions d'une variable complexe*, ENSTA ParisTech (2nd year)
- *Equations intégrales*, ENSTA ParisTech (3rd year) and Master "Modeling and Simulation" (M2)
- *Méthodes asymptotiques hautes fréquences pour les équations d'ondes - course notes*, ENSTA ParisTech (3rd year) and Master "Modeling and Simulation" (M2)

Eric Lunéville

- *Introduction au Calcul Scientifique*, ENSTA ParisTech (2nd year).
- *Programmation scientifique et simulation numérique*, ENSTA ParisTech (2nd year).
- *Propagation dans les guides d'ondes*, ENSTA ParisTech (3rd year) and Master "Modeling and Simulation" (M2)

Jean-François Mercier

- *Outils élémentaires d'analyse pour les équations aux dérivées partielles*, ENSTA ParisTech (1st year)
- *Fonctions d'une variable complexe*, ENSTA ParisTech, ENSTA ParisTech (1st year)
- *Théorie spectrale des opérateurs autoadjoints et application aux guides optiques*, ENSTA ParisTech (2nd year)

Axel Modave

- *Finite Element Methods*, ENSTA ParisTech (2nd year)
- *High Performance Scientific Computing*, ENSTA ParisTech (2rd year)
- *Parallel Scientific Computing*, ENSTA ParisTech (3rd year), and Master "Analysis, Modelling, Simulation" (M2)
- *Mathematical Models and their Discretisation in Electromagnetism*, ENSTA ParisTech (3rd year), and Master "Analysis, Modelling, Simulation" (M2)

10.2.2. Supervision

PhD: Luca Desiderio, "H-matrix based Solvers for 3D Elastodynamic Boundary Integral Equations", January 2017, Stéphanie Chaillat and Patrick Ciarlet

PhD: Laure Pesudo , "Modélisation de la réponse ultrasonore de défauts de type fissure par méthode BEM et couplage à un modèle de propagation - Application à la simulation des contrôle non destructifs", October 2017, Marc Bonnet

PhD in progress : Zouhair Adnani , "Modélisation numérique tridimensionnelle des effets de site en interaction sol-structure par une méthode adaptée aux problèmes sismiques de très grande taille", October 2014, Marc Bonnet and Stéphanie Chaillat

PhD in progress : Antoine Bensalah, "Une approche nouvelle de la modélisation mathématique et numérique en aéroacoustique par les équations de Goldstein et applications en aéronautique", October 2014, Patrick Joly and Jean-François Mercier

PhD in progress : Antoine Bera, "Conception de perturbations invisibles pour les ondes électromagnétiques ou acoustiques", October 2016, Anne-Sophie Bonnet-Ben Dhia and Lucas Chesnel

PhD in progress :Léandre Giret, "Development of a domain decomposition method on non-conforming meshes: application to the modeling of a Reactivity-Initiated Accident (RIA) in a Pressurized Water Reactor (PWR)", October 2014, Patrick Ciarlet

PhD in progress :Sandrine Paolantoni, "Analyse spectrale et simulation numérique de la diffraction électromagnétique par des métamatériaux", October 2016, Christophe Hazard and Boris Gralak

PhD in progress : Arnaud Recoquillay, "Identification de défauts dans un guide d'ondes en régime temporel", October 2014, Laurent Bourgeois

PhD in progress : Yohanes Tjandrawidjaja, "Modélisation de la propagation d'ondes guidées et de leur interaction avec des défauts localisés dans une plaque élastique anisotrope pour des applications en SHM", October 2016, Anne-Sophie Bonnet-Ben Dhia and Sonia Fliss

PhD in progress : Emile Parolin, "Non overlapping domain decomposition methods with non local transmission conditions for electromagnetic wave propagation", October 2017, PPatrick Joly and Xavier Claeys

PhD in progress : Clément Beneteau, "Asymptotic analysis of time harmonic Maxwell equations in presence of metamaterials", October 2017, Sonia Fliss and Xavier Claeys

PhD in progress : Hajer Methenni, "Mathematical modelling and numerical method for the simulation of ultrasound structural health monitoring of composite plates ", October 2017, Sonia Fliss and Sébastien Impériale

PhD in progress : Damien Mavaleix, ""Modeling of the fluid-structure interaction resulting from a remote underwater explosion", December 2017, Marc Bonnet and Stéphanie Chaillat

PhD in progress : Yacine Abourrig, "Boundary element method for modeling electromagnetic non-destructive testing: perturbative techniques for efficient and accurate parametric studies involving multiple simulations", October 2017, Marc Bonnet and Edouard Demaldent

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] L. DESIDERIO.*H-matrix based Solver for 3D Elastodynamics Boundary Integral Equations*, Université Paris-Saclay, January 2017, https://pastel.archives-ouvertes.fr/tel-01573534.

[2] L. PESUDO.*A hybrid strategy combining the integral equation method and the ray tracing method for high frequency diffraction involved in ultrasonic non destructive testing*, Université Paris-Saclay, October 2017, https://pastel.archives-ouvertes.fr/tel-01680366.

Articles in International Peer-Reviewed Journal

[3] M. BONNET.*A modified volume integral equation for anisotropic elastic or conducting inhomogeneities. Unconditional solvability by Neumann series*, in "Journal of Integral Equations and Applications", 2017, vol. 29, p. 271-295, https://hal.archives-ouvertes.fr/hal-01417944.

[4] M. BONNET, R. CORNAGGIA.*Higher order topological derivatives for three-dimensional anisotropic elasticity*, in "ESAIM: Mathematical Modelling and Numerical Analysis", 2017, n° 51, p. 2069-2092 [*DOI :* 10.1051/M2AN/2017015], https://hal.archives-ouvertes.fr/hal-01499498.

[5] A.-S. BONNET-BEN DHIA, C. CARVALHO, P. CIARLET.*Mesh requirements for the finite element approximation of problems with sign-changing coefficients*, in "Numerische Mathematik", 2017 [*DOI :* 10.1007/S00211-017-0923-5], https://hal.archives-ouvertes.fr/hal-01335153.

[6] A.-S. BONNET-BEN DHIA, L. CHESNEL, S. NAZAROV.*Perfect transmission invisibility for waveguides with sound hard walls*, in "Journal de Mathématiques Pures et Appliquées", August 2017, https://hal.archives-ouvertes.fr/hal-01371163.

[7] L. BOURGEOIS.*Quantification of the unique continuation property for the heat equation*, in "Mathematical Control and Related Fields", July 2017, vol. 7, n° 3, p. 347 - 367 [*DOI :* 10.3934/MCRF.2017012], https://hal.inria.fr/hal-01648045.

[8] L. BOURGEOIS, J. DARDÉ.*The "exterior approach" applied to the inverse obstacle problem for the heat equation*, in "SIAM Journal on Numerical Analysis", July 2017, vol. 55, n° 4, https://arxiv.org/abs/1609.05682 [*DOI :* 10.1137/16M1093872], https://hal.archives-ouvertes.fr/hal-01366850.

[9] E. BÉCACHE, P. JOLY, M. KACHANOVSKA.*Stable Perfectly Matched Layers for a Cold Plasma in a Strong Background Magnetic Field* , in "Journal of Computational Physics", July 2017, vol. 341, p. 76-101 [*DOI :* 10.1016/J.JCP.2017.03.051], https://hal.archives-ouvertes.fr/hal-01397581.

[10] E. BÉCACHE, M. KACHANOVSKA.*Stable perfectly matched layers for a class of anisotropic dispersive models. Part I: Necessary and sufficient conditions of stability: Extended Version*, in "ESAIM: Mathematical Modelling and Numerical Analysis", 2017, https://hal.archives-ouvertes.fr/hal-01356811.

[11] C. CARVALHO, L. CHESNEL, P. CIARLET.*Eigenvalue problems with sign-changing coefficients*, in "Comptes Rendus Mathématique", June 2017, vol. 355, n^{o} 6, p. 671 - 675 [*DOI :* 10.1016/J.CRMA.2017.05.002], https://hal.archives-ouvertes.fr/hal-01394856.

[12] M. CASSIER, C. HAZARD, P. JOLY.*Spectral theory for Maxwell's equations at the interface of a metamaterial. Part I: Generalized Fourier transform*, in "Communications in Partial Differential Equations", October 2017, vol. 42, n^{o} 11, p. 1707-1748, https://arxiv.org/abs/1610.03021 , https://hal-ensta.archives-ouvertes.fr/hal-01379118.

[13] M. CASSIER, P. JOLY, M. KACHANOVSKA.*Mathematical models for dispersive electromagnetic waves: An overview*, in "Computers and Mathematics with Applications", 2017, vol. 74, n^{o} 11, p. 2792-2830 [*DOI :* 10.1016/J.CAMWA.2017.07.025], https://hal.archives-ouvertes.fr/hal-01647185.

[14] S. CHAILLAT, M. DARBAS, F. LE LOUËR.*Fast iterative boundary element methods for high-frequency scattering problems in 3D elastodynamics*, in "Journal of Computational Physics", 2017, vol. 341, p. 429 - 446 [*DOI :* 10.1016/J.JCP.2017.04.020], https://hal.archives-ouvertes.fr/hal-01523020.

[15] S. CHAILLAT, L. DESIDERIO, P. CIARLET.*Theory and implementation of $\mathcal{H}$-matrix based iterative and direct solvers for Helmholtz and elastodynamic oscillatory kernels*, in "Journal of Computational Physics", September 2017, https://arxiv.org/abs/1706.09384 , https://hal.archives-ouvertes.fr/hal-01543919.

[16] P. CIARLET, S. FLISS, C. STOHRER.*On the Approximation of Electromagnetic Fields by Edge Finite Elements. Part 2: A Heterogeneous Multiscale Method for Maxwell's equations*, in "Computers and Mathematics with Applications", May 2017, vol. 73, n^{o} 9, p. 1900-1919 [*DOI :* 10.1016/J.CAMWA.2017.02.043], https://hal.archives-ouvertes.fr/hal-01364782.

[17] B. DELOURME, S. FLISS, P. JOLY, E. VASILEVSKAYA.*Trapped modes in thin and infinite ladder like domains. Part 1 : existence results*, in "Asymptotic Analysis", June 2017, vol. 103(3), n^{o} 103-134, https://arxiv.org/abs/1709.06345 , https://hal.inria.fr/hal-01287127.

[18] E. MARENIC, D. BRANCHERIE, M. BONNET.*Application of asymptotic analysis to the two-scale modeling of small defects in mechanical structures*, in "International Journal of Solids and Structures", 2017, vol. 128, p. 199-209, https://hal.archives-ouvertes.fr/hal-01588407.

[19] J.-F. MERCIER, B. LOMBARD.*A two-way model for nonlinear acoustic waves in a non-uniform lattice of Helmholtz resonators*, in "Wave Motion", 2017, vol. 72, p. 260-275, https://arxiv.org/abs/1609.03841 [*DOI :* 10.1016/J.WAVEMOTI.2017.04.004], https://hal.archives-ouvertes.fr/hal-01312801.

[20] A. MODAVE, A. ATLE, J. CHAN, T. WARBURTON.*A GPU-accelerated nodal discontinuous Galerkin method with high-order absorbing boundary conditions and corner/edge compatibility*, in "International Journal for Numerical Methods in Engineering", December 2017, vol. 112, n^{o} 11, p. 1659-1686, https://arxiv.org/abs/1610.05023 [*DOI :* 10.1002/NME.5576], https://hal.archives-ouvertes.fr/hal-01383074.

[21] A. MODAVE, J. LAMBRECHTS, C. GEUZAINE.*Perfectly matched layers for convex truncated domains with discontinuous Galerkin time domain simulations*, in "Computers and Mathematics with Applications",

February 2017, vol. 73, n^o^ 4, p. 684-700 [*DOI :* 10.1016/J.CAMWA.2016.12.027], https://hal.archives-ouvertes.fr/hal-01378501.

[22] A. OURIR, A. MAUREL, S. FÉLIX, J. F. MERCIER, M. FINK.*Manipulating light at subwavelength scale by exploiting defect-guided spoof plasmon modes*, in "Physical Review B : Condensed matter and materials physics", 2017, vol. 96, n^o^ 12 [*DOI :* 10.1103/PHYSREVB.96.125133], https://hal.inria.fr/hal-01626376.

Scientific Books (or Scientific Book chapters)

[23] F. ASSOUS, P. CIARLET, S. LABRUNIE.*Mathematical foundations of computational electromagnetism*, Springer, 2017, https://hal.archives-ouvertes.fr/hal-01596575.

[24] G. COHEN, S. PERNET.*Finite Element and Discontinuous Galerkin Methods for Transient Wave Equations*, Springer, 2017 [*DOI :* 10.1007/978-94-017-7761-2], https://hal.archives-ouvertes.fr/hal-01253166.

[25] J. LEBLOND, D. PONOMAREV.*On some extremal problems for analytic functions with constraints on real or imaginary parts*, in "Advances in Complex Analysis and Operator Theory: Festschrift in Honor of Daniel Alpay's 60th Birthday", F. COLOMBO, I. SABADINI, D. C. STRUPPA, M. B. VAJIAC (editors), Birkhauser, 2017, p. 219-236 [*DOI :* 10.1007/978-3-319-62362-7_8], https://hal.archives-ouvertes.fr/hal-01623281.

Other Publications

[26] A.-S. BONNET-BEN DHIA, L. CHESNEL, V. PAGNEUX.*Trapped modes and reflectionless modes as eigenfunctions of the same spectral problem*, January 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01692297.

[27] A.-S. BONNET-BEN DHIA, S. FLISS, A. TONNOIR.*The halfspace Matching Method : a new method to solve scattering problem in infinite media*, July 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01561339.

[28] P. CIARLET, C. F. DUNKL, S. A. SAUTER.*A Family of Crouzeix-Raviart Finite Elements in 3D*, February 2017, working paper or preprint, https://hal.inria.fr/hal-01488172.

[29] P. CIARLET, L. GIRET, E. JAMELOT, F. D. KPADONOU.*Numerical analysis of the mixed finite element method for the neutron diffusion eigenproblem with heterogeneous coefficients*, December 2017, working paper or preprint, https://hal.inria.fr/hal-01566179.

[30] P. CIARLET, E. JAMELOT, F. D. KPADONOU.*Domain decomposition methods for the diffusion equation with low-regularity solution*, July 2017, working paper or preprint, https://hal.inria.fr/hal-01349385.

[31] P. CIARLET, M. VOHRALÍK.*Localization of global norms and robust a posteriori error control for transmission problems with sign-changing coefficients*, August 2017, working paper or preprint, https://hal.inria.fr/hal-01148476.

[32] C. HAZARD, S. PAOLANTONI.*Spectral analysis of polygonal cavities containing a negative-index material*, October 2017, working paper or preprint, https://hal-ensta.archives-ouvertes.fr/hal-01626868.

[33] M. KACHANOVSKA.*Stable perfectly matched layers for a class of anisotropic dispersive models. Part II: Energy estimates*, March 2017, This work was supported by a public grant as part of the Investissement

d'avenir project, reference ANR-11-LABX-0056-LMH, LabEx LMH, as well as a co- financing program PRESTIG, https://hal.inria.fr/hal-01419682.

[34] J. F. MERCIER, C. MIETKA, F. MILLOT, V. PAGNEUX.*Acoustic propagation in a vortical homentropic flow*, December 2017, working paper or preprint, https://hal.inria.fr/hal-01663949.

Team RANDOPT

Randomized Optimization

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions).

RESEARCH CENTER
Saclay - Île-de-France

THEME
Optimization, machine learning and statistical methods

Table of contents

Team RANDOPT

Creation of the Team: 2016 December 01

Keywords:

Computer Science and Digital Science:

A6. - Modeling, simulation and control
A6.2. - Scientific Computing, Numerical Analysis & Optimization
A6.2.2. - Numerical probability
A6.2.3. - Probabilistic methods
A6.2.4. - Statistical methods
A6.2.6. - Optimization
A8.2.2. - Evolutionary algorithms
A8.9. - Performance evaluation

Other Research Topics and Application Domains:

B4. - Energy

1. Personnel

Research Scientists

Anne Auger [Team leader, Inria, Researcher]
Dimo Brockhoff [Inria, Researcher]
Nikolaus Hansen [Inria, Senior Researcher, HDR]

Technical Staff

Asma Atamna [Inria, until Jan 2017]
Umut Batu [Inria, from Jul 2017]
Dejan Tusar [Inria, from Nov 2017]

PhD Students

Cheikh Saliou Toure [Inria, from Oct 2017]
Konstantinos Varelas [Thales, from Dec 2017]

Post-Doctoral Fellow

Asma Atamna [Inria, from Feb 2017]

Administrative Assistants

Natalia Alves [Inria]
Olga Mwana Mobulakani [Inria]

2. Overall Objectives

2.1. Scientific Context

Important problems in various scientific domains like biology, physics, medicine or in industry critically rely on the resolution of difficult numerical optimization problems. Often those problems depend on noisy data or are the outcome of complex numerical simulations such that derivatives are not available or not useful and the function is seen as a black-box.

Many of those optimization problems are in essence *multiobjective*—one needs to optimize simultaneously several conflicting objectives like minimizing the cost of an energy network and maximizing its reliability—and most of the *challenging* black-box problems are *non-convex, non-smooth* and combine difficulties related to ill-conditioning, non-separability, and ruggedness (a term that characterizes functions that can be non-smooth but also noisy or multi-modal). Additionally, the objective function can be expensive to evaluate—a single function evaluation might take several minutes to hours (it can involve for instance a CFD simulation).

In this context, the use of randomness combined with proper adaptive mechanisms has proven to be one key component for the design of robust global numerical optimization algorithms.
The field of adaptive stochastic optimization algorithms has witnessed some important progress over the past 15 years. On the one hand, subdomains like medium-scale unconstrained optimization may be considered as "solved" (particularly, the CMA-ES algorithm, an instance of *Evolution Strategy* (ES) algorithms, stands out as state-of-the-art method) and considerably better standards have been established in the way benchmarking and experimentation are performed. On the other hand, multiobjective population-based stochastic algorithms became the method of choice to address multiobjective problems when a set of some best possible compromises is sought for. In all cases, the resulting algorithms have been naturally transferred to industry (the CMA-ES algorithm is now regularly used in companies such as Bosch, Total, ALSTOM, ...) or to other academic domains where difficult problems need to be solved such as physics, biology [28], geoscience [23], or robotics [25].

Very recently, ES algorithms attracted quite some attention in Machine Learning with the OpenAI article *Evolution Strategies as a Scalable Alternative to Reinforcement Learning*. It is shown that the training time for difficult reinforcement learning benchmarks could be reduced from 1 day (with standard RL approaches) to 1 hour using ES [27]. A few years ago, another impressive application of CMA-ES, how "Computer Sim Teaches Itself To Walk Upright" (published at the conference SIGGRAPH Asia 2013) was presented in the press in the UK.

Several of those important advances around adaptive stochastic optimization algorithms are relying to a great extent on works initiated or achieved by the founding members of RandOpt particularly related to the CMA-ES algorithm and to the Comparing Continuous Optimizer (COCO) platform (see Section on Software and Platform).
Yet, the field of adaptive stochastic algorithms for black-box optimization is relatively young compared to the "classical optimization" field that includes convex and gradient-based optimization. For instance, the state-of-the art algorithms for unconstrained gradient based optimization like quasi-newton methods (e.g. the BFGS method) date from the 1970s [20] while the stochastic derivative-free counterpart, CMA-ES dates from the early 2000s [21]. Consequently, in some subdomains with *important practical demands*, not even the most fundamental and basic questions are answered:

- This is the case of *constrained* optimization where one needs to find a solution $x^* \in \mathbb{R}^n$ minimizing a numerical function $\min_{x \in \mathbb{R}^n} f(x)$ while respecting a number of constraints m typically formulated as $g_i(x^*) \leq 0$ for $i = 1, ..., m$. Only recently, the fundamental requirement of linear convergence, as in the unconstrained case, has been clearly stated [13].
- In multiobjective optimization, most of the research so far has been focusing on *how to select candidate solutions from one iteration to the next one*. The difficult question of how to *generate* effectively new solutions is not yet answered in a proper way and we know today that simply applying operators from single-objective optimization may not be effective with the current best selection strategies. As a comparison, in the single-objective case, the question of selection of candidate solutions was already solved in the 1980s and 15 more years where needed to solve the trickier question of an effective adaptive strategy to generate new solutions.
- With the current demand to solve larger and larger optimization problems (e.g. in the domain of deep learning), optimization algorithms that scale linearly with the problem dimension are nowadays needed. Only recently, first proposals of how to reduce the quadratic scaling of CMA-ES have been made without a clear view of what can be achieved in the best case *in practice*. These later

variants apply to medium scale optimization with thousands of variables. The question of designing randomized algorithms capable to handle efficiently problems with one or two orders of magnitude more variables is still largely open.

- For expensive optimization, standard methods are so called Bayesian optimization algorithms based on Gaussian processes. In this domain, no standard implementations exist and performance across implementations can vary significantly as particularly different algorithm components are used. For instance, there is no common agreement on which initial design to use, which bandwidth for the Gaussian kernel to take, and which strategy to follow to optimize the expected improvement.

2.2. Overall Objectives

In the context of black-box numerical optimization previously described, the main ambition of the RandOpt team is **to design and implement novel methods in subdomains with a strong practical demand**. Those methods should become future standards that allow to solve important challenging applications in industry or academia. For this, we believe that (i) **theory** can greatly help for **algorithm design**; (ii) the **development and implementation of proper scientific experimentation methodology** is crucial and (iii) it is decisive to provide **parameter-less implementations** of the methods through open-source software packages. This shapes four main scientific goals for our proposed team:

1. **develop novel theoretical frameworks** for guiding (a) the design of novel methods and (b) their analysis, allowing to
2. provide **proofs of key features** of stochastic adaptive algorithms including the state-of-the-art method CMA-ES: linear convergence and learning of second order information.
3. develop **novel stochastic numerical black-box algorithms** following a **principled design** in domains with a strong practical need for much better methods namely **constrained, multiobjective, large-scale and expensive optimization**. Implement the methods such that they are easy to use. And finally, to
4. **set new standards in scientific experimentation, performance assessment and benchmarking** both for optimization on continuous or combinatorial search spaces. This should allow in particular to advance the state of **reproducibility of results of scientific papers** in optimization.

All the above relate to our objectives with respect to dissemination and transfer:

1. develop software packages that people can directly use to solve their problems. This means having carefully thought out interfaces, generically applicable setting of parameters and termination conditions, proper treatment of numerical errors, catching properly various exceptions, etc.;
2. have direct collaborations with industrials;
3. publish our results both in applied mathematics and computer science bridging the gap between very often disjoint communities.

3. Research Program

3.1. Developing Novel Theoretical Frameworks for Analyzing and Designing Adaptive Stochastic Algorithms

The lines of research of the RandOpt team are organized along four axis namely developing novel theoretical framework, developing novel algorithms, setting novel standards in scientific experimentation and benchmarking and applications.

Stochastic black-box algorithms typically optimize **non-convex, non-smooth functions**. This is possible because the algorithms rely on weak mathematical properties of the underlying functions: not only derivatives (gradients) are not exploited, but often the methods are so-called comparison-based which means that the algorithm will only rely on the ranking of the candidate solutions' function values. This renders those methods more robust as they are invariant to strictly increasing transformations of the objective function but at the same time the theoretical analysis becomes more difficult as **we cannot exploit a well defined framework using (strong) properties of the function like convexity or smoothness**.

Additionally, adaptive stochastic optimization algorithms typically have a **complex state space** which encodes the parameters of a probability distribution (e.g. mean and covariance matrix of a Gaussian vector) and other state vectors. This state-space is a **manifold**. While the algorithms are Markov chains, the complexity of the state-space makes that **standard Markov chain theory tools do not directly apply**. The same holds with tools stemming from stochastic approximation theory or Ordinary Differential Equation (ODE) theory where it is usually assumed that the underlying ODE (obtained by proper averaging and limit for learning rate to zero) has its critical points inside the search space. In contrast, in the cases we are interested, the **critical points of the ODEs are at the boundary of the domain**.

Last, since we aim at developing theory that one the one hand allows to analyze the main properties of state-of-the-art methods and on the other hand is useful for algorithm design, we need to be careful to not use simplifications that would allow a proof to be done but would not capture the important properties of the algorithms. With that respect one tricky point is to develop **theory that accounts for invariance properties**. To face those specific challenges, we need to develop novel theoretical frameworks exploiting invariance properties and accounting for peculiar state-spaces. Those frameworks should allow to analyze one of the core properties of adaptive stochastic methods, namely **linear convergence** on the widest possible class of functions.

We are planning on approaching the question of linear convergence from three different complementary angles, using three different frameworks:

- the Markov chain framework where the convergence derives from the analysis of the stability of a normalized Markov chain existing on scaling-invariant functions for translation and scale-invariant algorithms [15]. This framework allows for a fine analysis where the exact convergence rate can be given as an implicit function of the invariant measure of the normalized Markov chain. Yet it requires the objective function to be scaling-invariant. The stability analysis can be particularly tricky as the Markov chain that needs to be studied writes as $\Phi_{t+1} = F(\Phi_t, W_{t+1})$ where $\{W_t : t > 0\}$ are independent identically distributed and F is typically discontinuous because the algorithms studied are comparison-based. This implies that practical tools for analyzing a standard property like irreducibility, that rely on investigating the stability of underlying deterministic control models [26], cannot be used. Additionally, the construction of a drift to prove ergodicity is particularly delicate when the state space includes a (normalized) covariance matrix as it is the case for analyzing the CMA-ES algorithm.
- The stochastic approximation or ODE framework. Those are standard techniques to prove the convergence of stochastic algorithms when an algorithm can be expressed as a stochastic approximation of the solution of a mean field ODE [16], [17], [24]. What is specific and induces difficulties for the algorithms we aim at analyzing is the **non-standard state-space** since the ODE variables correspond to the state-variables of the algorithm (e.g. $\mathbb{R}^n \times \mathbb{R}_{>0}$ for step-size adaptive algorithms, $\mathbb{R}^n \times \mathbb{R}_{>0} \times S^n_{++}$ where S^n_{++} denotes the set of positive definite matrices if a covariance matrix is additionally adapted). Consequently, the ODE can have many critical points at the boundary of its definition domain (e.g. all points corresponding to $\sigma_t = 0$ are critical points of the ODE) which is not typical. Also we aim at proving **linear convergence**, for that it is crucial that the learning rate does not decrease to zero which is non-standard in ODE method.
- The direct framework where we construct a global Lyapunov function for the original algorithm from which we deduce bounds on the hitting time to reach an ϵ-ball of the optimum. For this framework as for the ODE framework, we expect that the class of functions where we can prove linear convergence

are composite of $g \circ f$ where f is differentiable and $g : \mathrm{Im}(f) \to \mathbb{R}$ is strictly increasing and that we can show convergence to a local minimum.

We expect those frameworks to be complementary in the sense that the assumptions required are different. Typically, the ODE framework should allow for proofs under the assumptions that learning rates are small enough while it is not needed for the Markov chain framework. Hence this latter framework captures better the real dynamics of the algorithm, yet under the assumption of scaling-invariance of the objective functions. By studying the different frameworks in parallel, we expect to gain synergies and possibly understand what is the most promising approach for solving the holy grail question of the linear convergence of CMA-ES.

3.2. Developing Novel Adaptive Stochastic Algorithms

We are planning on developing novel algorithms in the subdomains with strong practical demand for better methods of constrained, multiobjective, large-scale and expensive optimization.

3.2.1. Constrained optimization

Many (real-world) optimization problems have constraints related to technical feasibility, cost, etc. Constraints are classically handled in the black-box setting either via rejection of solutions violating the constraints—which can be quite costly and even lead to quasi-infinite loops—or by penalization with respect to the distance to the feasible domain (if this information can be extracted) or with respect to the constraint function value [18]. However, the penalization coefficient is a sensitive parameter that needs to be adapted in order to achieve a robust and general method [19]. Yet, **the question of how to handle properly constraints is largely unsolved**. The latest constraints handling for CMA-ES is an ad-hoc technique driven by many heuristics [19]. Also, it is particularly only recently that it was pointed out that **linear convergence properties should be preserved** when addressing constraint problems [13].

Promising approaches though, rely on using augmented Lagrangians [13], [14]. The augmented Lagrangian, here, is the objective function optimized by the algorithm. Yet, it depends on coefficients that are adapted online. The adaptation of those coefficients is the difficult part: the algorithm should be stable and the adaptation efficient. We believe that the theoretical frameworks developed (particularly the Markov chain framework) will be useful to understand how to design the adaptation mechanisms. Additionally, the question of invariance will also be at the core of the design of the methods: augmented Lagrangian approaches break the invariance to monotonic transformation of the objective functions, yet understanding the maximal invariance that can be achieved seems to be an important step towards understanding what adaptation rules should satisfy.

3.2.2. Large-scale Optimization

In the large-scale setting, we are interested to optimize problems with the order of 10^3 to 10^4 variables. For one to two orders of magnitude more variables, we will talk about a “very large-scale" setting.

In this context, algorithms with a quadratic scaling (internal and in terms of number of function evaluations needed to optimize the problem) cannot be afforded. In CMA-ES-type algorithms, we typically need to restrict the model of the covariance matrix to have only a linear number of parameters to learn such that the algorithms scale linearly. The main challenge is thus to have rich enough models for which we can efficiently design proper adaptation mechanisms. Some first large-scale variants of CMA-ES have been derived. They include the online adaptation of the complexity of the model [12], [11]. Yet there are still open problems related to being able to learn both short and long axes in the models.

Another direction, we want to pursue, is exploring the use of large-scale variants of CMA-ES to solve reinforcement learning problems [27].

Last, we are interested to investigate the very-large-scale setting. One approach consist in doing optimization in subspaces. This entails the efficient identification of relevant spaces and the restriction of the optimization to those subspaces.

3.2.3. Multiobjective Optimization

Multiobjective optimization, i.e., the simultaneous optimization of multiple objective functions, differs from single-objective optimization in particular in its optimization goal. Instead of aiming at converging to the solution with the best possible function value, in multiobjective optimization, a set of solutions[0] is sought. This set, called Pareto-set, contains all trade-off solutions in the sense of Pareto-optimality—no solution exists that is better in *all* objectives than a Pareto-optimal one. Because converging towards a set differs from converging to a single solution, it is no surprise that we might lose many good convergence properties if we directly apply search operators from single-objective methods. However, this is what has typically been done so far in the literature. Indeed, most of the research in stochastic algorithms for multiobjective optimization focused instead on the so called selection part, that decides which solutions should be kept during the optimization—a question that can be considered as solved for many years in the case of single-objective stochastic adaptive methods.

We therefore aim at rethinking search operators and adaptive mechanisms to improve existing methods. We expect that we can obtain orders of magnitude better convergence rates for certain problem types if we choose the right search operators. We typically see two angles of attack: On the one hand, we will study methods based on scalarizing functions that transform the multiobjective problem into a set of single-objective problems. Those single-objective problems can then be solved with state-of-the-art single-objective algorithms. Classical methods for multiobjective optimization fall into this category, but they all solve multiple single-objective problems subsequently (from scratch) instead of dynamically changing the scalarizing function during the search. On the other hand, we will improve on currently available population-based methods such as the first multiobjective versions of the CMA-ES. Here, research is needed on an even more fundamental level such as trying to understand success probabilities observed during an optimization run or how we can introduce non-elitist selection (the state of the art in single-objective stochastic adaptive algorithms) to increase robustness regarding noisy evaluations or multi-modality. The challenge here, compared to single-objective algorithms, is that the quality of a solution is not anymore independent from other sampled solutions, but can potentially depend on all known solutions (in the case of three or more objective functions), resulting in a more noisy evaluation as the relatively simple function-value-based ranking within single-objective optimizers.

3.2.4. Expensive Optimization

In the so-called expensive optimization scenario, a single function evaluation might take several minutes or even hours in a practical setting. Hence, the available budget in terms of number of function evaluation calls to find a solution is very limited in practice. To tackle such expensive optimization problems, it is needed to exploit the first few function evaluations in the best way. To this end, typical methods couple the learning of a surrogate (or meta-model) of the expensive objective function with traditional optimization algorithms.

In the context of expensive optimization and CMA-ES, which usually shows its full potential when the number n of variables is not too small (say larger than 3) and if the number of available function evaluations is about $100n$ or larger, several research directions emerge. The two main possibilities to integrate meta-models into the search with CMA-ES type algorithms are (i) the successive injection of the minimum of a learned meta-model at each time step into the learning of CMA-ES's covariance matrix and (ii) the use of a meta-model to predict the internal ranking of solutions. While for the latter, first results exist, the former idea is entirely unexplored for now. In both cases, a fundamental question is which type of meta-model (linear, quadratic, Gaussian Process, ...) is the best choice for a given number of function evaluations (as low as one or two function evaluations) and at which time the type of the meta-model shall be switched.

3.3. Setting novel standards in scientific experimentation and benchmarking

Numerical experimentation is needed as a complement to theory to test novel ideas, hypotheses, the stability of an algorithm, and/or to obtain quantitative estimates. Optimally, theory and experimentation go hand in hand, jointly guiding the understanding of the mechanisms underlying optimization algorithms. Though performing

[0]Often, this set forms a manifold of dimension one smaller than the number of objectives.

numerical experimentation on optimization algorithms is crucial and a common task, it is non-trivial and easy to fall in (common) pitfalls as stated by J. N. Hooker in his seminal paper [22].

In the RandOpt team we aim at raising the standards for both scientific experimentation and benchmarking.

On the experimentation aspect, we are convinced that there is common ground over how scientific experimentation should be done across many (sub-)domains of optimization, in particular with respect to the visualization of results, testing extreme scenarios (parameter settings, initial conditions, etc.), how to conduct understandable and small experiments, how to account for invariance properties, performing scaling up experiments and so forth. We therefore want to formalize and generalize these ideas in order to make them known to the entire optimization community with the final aim that they become standards for experimental research.

Extensive numerical benchmarking, on the other hand, is a compulsory task for evaluating and comparing the performance of algorithms. It puts algorithms to a standardized test and allows to make recommendations which algorithms should be used preferably in practice. To ease this part of optimization research, we have been developing the Comparing Continuous Optimizers platform (COCO) since 2007 (see also the software section below) which allows to automatize the tedious task of benchmarking. It is a game changer in the sense that the freed time can now be spent on the scientific part of algorithm design (instead of implementing the experiments, visualization, statistical tests, etc.) and it opened novel perspectives in algorithm testing. COCO implements a thorough, well-documented methodology that is based on the above mentioned general principles for scientific experimentation.

Also due to the freely available data from 200+ algorithms benchmarked with the platform, COCO became a quasi-standard for single-objective, noiseless optimization benchmarking. It is therefore natural to extend the reach of COCO towards other subdomains (particularly constrained optimization, many-objective optimization) which can benefit greatly from an automated benchmarking methodology and standardized tests without (much) effort. This entails particularly the design of novel test suites and rethinking the methodology for measuring performance and more generally evaluating the algorithms. Particularly challenging is the design of scalable non-trivial testbeds for constrained optimization where one can still control where the solutions lies. Other optimization problem types, we are targeting are expensive problems (and the Bayesian optimization community in particular, see our AESOP project), optimization problems in machine learning (for example parameter tuning in reinforcement learning), and the collection of real-world problems from industry.

Another aspect of our future research on benchmarking is to investigate the large amounts of benchmarking data, we collected with COCO during the years. Extracting information about the influence of algorithms on the best performing portfolio, clustering algorithms of similar performance, or the automated detection of anomalies in terms of good/bad behavior of algorithms on a subset of the functions or dimensions are some of the ideas here.

Last, we want to expand the focus of COCO from automatized (large) benchmarking experiments towards everyday experimentation, for example by allowing the user to visually investigate algorithm internals on the fly or by simplifying the set up of algorithm parameter influence studies.

4. Application Domains

4.1. Applications

Applications of black-box algorithms occur in various domains. Industry but also researchers in other academic domains have therefore a great need to apply black-box algorithms on a daily basis. We see this as a great source of motivation to design better methods. Applications not only allow us to backup our methods and understand what are the relevant features to solve a real-world problem but also help to identify novel difficulties or set priorities in terms of algorithm design.

We are currently dealing with concrete applications related to three industrial collaborations:

- With EDF R&D through the design and placement of bi-facial photovoltaic panels for the postdoc of Asma Atamna funded by the PGMO project NumBER.
- With Thales for the PhD thesis of Konstantinos Varelas (DGA-CIFRE thesis) related to applications in the defense domain.
- With Storengy, a subsidiary of Engie specialized in gas storage, for the PhD thesis of Cheikh Touré.

Another type of application we want to focus on comes from reinforcement learning. The problems addressed in [27] seem to be particularly suited for large-scale variants of CMA-ES.

When dealing with single applications, the results observed are difficult to generalize: typically not many methods are tested on a single application as tests are often time consuming and performed in restrictive settings. Yet, if one circumvent the problem of confidentiality of data and of criticality for companies to publish their applications, real-world problems could become benchmarks as any other analytical function. This would allow to test wider ranges of methods on the problems and to find out whether analytical benchmarks properly capture real-world problem difficulties. We will thus seek to incorporate real-world problems within our COCO platform. This is a recurrent demand by researchers in optimization.

5. Highlights of the Year

5.1. Highlights of the Year

- A Auger has been (re)-elected member of the ACM-SIGEVO executive board.

6. New Results

6.1. Theory

Participants: Anne Auger, Nikolaus Hansen.

The paper “Information-Geometric Optimization Algorithms: A Unifying Picture via Invariance Principles" has finally been published in the JMLR journal [3]. In this paper in collaboration with Yann Ollivier in particular, we lay the ground of stochastic optimization by means of information geometry. We provide a unified framework for stochastic optimization on arbitrary search spaces that allow to recover well-known algorithms on continuous or discrete search spaces and put them under the same umbrella of Information Geometric Optimization.

When analyzing the stability of Markov chains stemming from comparison-based stochastic algorithms, we are facing difficulties due to the fact that the Markov chains have the following form $\Phi_{t+1} = F(\Phi_t, U_{t+1})$ where $\{U_t : t \geq 0\}$ are i.i.d. random vectors and F is a discontinuous function. The discontinuity comes from the comparison-based property of the algorithms. If F were C^∞ or C^1 we could prove easily stability properties like irreducibility and show that compact are small sets by investigating the underlying control model and showing that it has globally attracting states where controllability conditions hold using results developed by Sean Meyn and co-authors.

In the paper [2], we found that we can actually generalize to a great extent the results by Meyn to the case where $\Phi_{t+1} = F(\Phi_t, \alpha(\Phi_t, U_{t+1}))$ where F is C^1 and α is discontinuous but such that $\alpha(x, U)$ admits a lower-semi continuity density. We have proposed verifiable conditions for the irreducibility and aperiodicity and shown that compact sets are small sets.

The development of evolution strategies has been greatly driven by so-called progress rate or quality gain analysis where simplification assumptions are made to obtain quantitative estimate of progress in one step and deduce from it how to set different parameters like recombination weights, learning rates ...

This theory while very useful often relied on approximations that were not always well appraised, justified or clearly stated. We have been in the past rigorously deriving different progress rate results and related them to bounds on convergence rates. We have investigated rigorously the quality gain (that is progress measured in terms of objective function) on general convex quadratic functions using weighted recombination. This allowed to derive the dependency of the convergence rate of evolution strategies with respect to the eigenspectrum of the Hessian matrix of convex-quadratic function as well as give hints on how to set learning rate [4] and [9].

6.2. Novel Constraint Handling

Participants: Asma Atamna, Anne Auger, Nikolaus Hansen.

In the context of constrained optimization, we have investigated to use augmented Lagrangian approaches to handle constraints. The originality of the approach is that the parameters of the augmented Lagrangian are adapted online. We have shown sufficient conditions for linear convergence of the ensuing methods with linear constraints [5]. Those sufficient conditions rely on finding a Markov chain candidate to be stable. This Markov chain derives from invariance properties of the algorithm. At the same time we have proposed an algorithm variant for the $(\mu/\mu, \lambda)$-CMA-ES and an arbitrary number of constraints.

In [10], we have investigated the linear convergence question on the point of view of invariance. We have analyzed the invariances of adaptive algorithms handling constraints with augmented Lagrangian: we have shown that invariance to monotonic transformation of the objective functions is lost but that a subclass of invariance can and should be preserved, namely affine transformation of the objective function and scaling of the constraint by a positive constant.

6.3. Benchmarking

Participants: Anne Auger, Dimo Brockhoff, Nikolaus Hansen, Umut Batu, Dejan Tusar.

In his thesis, Ouassim AitElHara has been investigating the benchmarking of algorithms in large dimensions [1]. In this context, the first steps for a testbed of the COCO platform in large dimension have been done. Particularly, the methodology for building a large-scale testbed has been defined: it consists in replacing the usual orthogonal transformation by block-diagonal orthogonal matrices multiplied to the left and to the right by permutation matrices. While still under testing, we expect to be able to release the large-scale testbed in the coming year.

The population size is one of the few parameters, a user is supposed to touch in the state-of-the-art optimizer CMA-ES. In [7], a new approach to also adapt the population size in CMA-ES is proposed and benchmarked on the bbob test suite of our COCO platform. The method is based on tracking the non-decrease of the median of the objective function values in each slot of S successive iterations to decide whether to increase or decrease or keep the population size in the next slot of S iterations. The experimental results show the efficiency of our approach on some multi-modal functions with adequate global structure.

Benchmarking budget-dependent algorithms (for which parameters might depend on the given budget of function evaluations) is typically done for a fixed (set of) budget(s). This, however, has the disadvantage that the reported function values at this budget are difficult to interpret. Furthermore, assessing performance in this way does not give any hints how an algorithm would behave for other budgets. Instead, we proposed in [8] a new way to do “Anytime Benchmarking of Budget-Dependent Algorithms” and implemented this functionality in our COCO platform. The idea is to run several experiments for varying budgets and report target-based runtimes in the form of empirical cumulative distribution functions (aka data profiles) as in the case of anytime algorithms.

6.4. Performance Assessment in Multiobjective Optimization

In the context of performance assessment in multiojective optimization, two contributions have been made in 2017. First, we proposed a new visualization method to quantitatively assess the performance of multiobjective optimizers (for 2-objective problems) in the form of average runtime attainment functions [6]. The main idea is to display, for each point in objective space, when (in terms of the average runtime) it has been attained or in other words when it has been dominated by the algorithm. Second, we continued our effort towards automated benchmarking via our COCO platform and described a generic test suite generator that can produce test suites like the previous `bbob-biobj` test suite for an arbitrary number of objectives[0].

6.5. Comparing Continous Optimizers Platform

Participants: Anne Auger, Dimo Brockhoff, Nikolaus Hansen, Umut Batu, Dejan Tusar.

Thanks to the ADT support for Dejan Tušar (since November, previously supported by ESA) and Umut Batu (since July), as well as due to an increased effort from the core development team, we could progress on several aspects regarding our Comparing Continuous Optimizers platform (COCO, https://github.com/numbbo/coco) in 2017.

Most notably, we provide the new functionality of data archives which allows to access the available data of 200+ algorithms much easier. We also made significant progress towards a first constrained test suite—in particular did we add logging support for constrained problems. The postprocessing module is finally python 3 compatible and zip files are supported as input files. The reference worst f-values-of-interest are exposed to the (multiobjective) solver, algorithms can now be displayed in the background, and simplified example experiment scripts (in python) are available (for both anytime, and budget-dependend algorithms, see also [8]). We also improved our continuous integration support, now using also CircleCI and AppVeyor in addition to Inria's Jenkins system. Version 2.0, released in January 2017, saw new functionality of reference algorithms for the multiobjective test suite, a new format of reference algorithms that allow to use any existing data set as reference, improved HTML output and navigation, the COCO version number being part of the plots now, and new regression tests for all provided test suites.

COCO facts for 2017

- 218 issues closed
- major release 2.0 in January plus three additional releases, version 2.2 planned for January 2018
- 10 new contributors outside the main development team
- 14 new algorithm data sets made public (total: 233)

Currently, we are working on an entire rewrite of the postprocessing (ADT COCOpost project of Umut Batu), an improved cocoex module for proposing test suites, functions, data loggers etc. in python (ADT COCOpysuites of Dejan Tušar), a first constrained test suite (in particular Asma Atamna via the PGMO project NumBER), and a large-scale test suite (part of Konstantinos Varelas' PhD thesis, based on the PhD work of Ouassim AitElHara).

Finally, we continued to use COCO also for teaching, in particular for the group project ("controle continue") of our Introduction to Optimization (about 40 Master students) and the Derivative-Free Optimization lectures at Université Paris-Sud (about 30 Master students).

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Grants and Contracts with Industry

- CIFRE-DGA with Thales, for the PhD of Konstantinos Varelas (2017—2020)

[0] Paper to be submitted to the Evolutionary Computation journal and to arXiv/HAL in January 2018.

- contract with Storengy to finance a part of the PhD of Cheikh Touré (2017—2020)

8. Partnerships and Cooperations

8.1. Regional Initiatives

- PGMO project “NumBER: Numerical Black Box Optimization for Energy Applications”, in collaboration with EDF, financing the postdoc of Asma Atamna, project length: 2 years (2016–2018), PI: Anne Auger
- PGMO project “AESOP: Algorithms Expensive Simulation-Basd Optimization Problems”, a project involving several researchers from CentraleSupelec, Ecole des Mines de St.-Etienne, INRA Toulouse, JSI (Slovenia), Safran, Ruhr-Universität Bochum (Germany), and TU Dortmund University (Germany), project length: 2 years (2017–2019), PI: Dimo Brockhoff

8.2. National Initiatives

8.2.1. ANR

- ANR project “NumBBO: Analysis, Improvement and Evaluation of Numerical Blackbox Optimizers”, with partners DOLPHIN team (till 2016), Ecole des Mines de St.-Etienne and TU Dortmund University (Germany), Anne Auger was PI of this project which had a total budget of 660kEUR (2012–2017)
- ANR project “Big Multiobjective Optimization (BigMO)”, Dimo Brockhoff participates in this project through the Inria team BONUS in Lille (2017–2020)

8.3. International Initiatives

8.3.1. Inria Associate Teams Not Involved in an Inria International Labs

8.3.1.1. s3-bbo

Title: Threefold Scalability in Any-objective Black-Box Optimization

International Partner (Institution - Laboratory - Researcher):

Shinshu (Japan) - Tanaka-Hernan-Akimoto Laboratory - Hernan Aguirre

Start year: 2015

See also: http://francejapan.gforge.inria.fr/doku.php?id=associateteam

This associate team brings together researchers from the TAO and Dolphin Inria teams with researchers from Shinshu university in Japan. Additionally, researchers from the University of Calais are external collaborators to the team. The common interest is on black-box single and multi-objective optimization with complementary expertises ranging from theoretical and fundamental aspects over algorithm design to solving industrial applications. The work that we want to pursue in the context of the associate team is focused on black-box optimization of problems with a large number of decision variables and one or several functions to evaluate solutions, employing distributed and parallel computing resources. The objective is to theoretically derive, analyze, design, and develop scalable black-box stochastic algorithms including evolutionary algorithms for large-scale optimization considering three different axes of scalability: (i) decision space, (ii) objective space, and (iii) availability of distributed and parallel computing resources.

We foresee that the associate team will make easier the collaboration already existing through a proposal funded by Japan and open-up a long term fruitful collaboration between Inria and Shinshu university. The collaboration will be through exchanging researchers and Ph.D. students and co-organization of workshops.

8.3.2. Inria International Partners

8.3.2.1. Declared Inria International Partners

- We are collaborating with Shinshu University and particularly Youhei Akimoto through our joint associate team.

8.3.2.2. Informal International Partners

- We are collaborating with Tea Tušar from the Josef-Stefan Institute in Ljubljana, Slovenia for extending and maintaining our COCO platform and on benchmarking in general.
- We are collaborating with Jun.-Prof. Tobias Glasmachers from the Ruhr-Universität Bochum in Germany on runtime analysis of adaptive stochastic algorithms.

8.4. International Research Visitors

8.4.1. Visits of International Scientists

- Filip Matzner from Charles University Prague (Czech Republic) - Visit of one month in November 2017 to work on Evolution Strategies for reinforcement learning and classification problems.
- Prof. Dr. Youhei Akimoto from Shinhu University (Japan) - Visit of one month in November 2017 to work on several projects related to theory and algorithm design for large-scale optimization.
- Dr. Alexandre Chotard from KTU (Sweden) - Visit of one month in November 2017 to work on adaptive MCMC.
- Dr. Tea Tušar from the Josef-Stefan Institute (Slovenia) - Visit of one week in November 2017 to work on our projects around (multiobjective) blackbox optimization benchmarking.

8.4.2. Visits to International Teams

8.4.2.1. Research Stays Abroad

- Anne Auger and Dimo Brockhoff visited Jun.-Prof. Tobias Glasmachers and Prof. Günter Rudolph in Dortmund from April 10 till April 14, 2017

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

- A. Auger, program chair of the PPSN 2018 conference Coimbra, Portugal

9.1.1.2. Member of the Organizing Committees

- Anne Auger, Dimo Brockhoff, Nikolaus Hansen, and Dejan Tušar, co-organizer of the ACM-GECCO-2017 workshop on Black Box Optimization Benchmarking, together with Tea Tušar
- Anne Auger, Dimo Brockhoff and Nikolaus Hansen, co-organizer of the ACM-GECCO-2018 workshop on Black Box Optimization Benchmarking, together with Julien Bect, Rodolphe Le Riche, Victor Picheny, and Tea Tušar

9.1.2. Scientific Events Selection

9.1.2.1. Chair of Conference Program Committees

- Anne Auger: theory track chair for the ACM-GECCO conference 2018, Kyoto, Japan
- Nikolaus Hansen: co-track chair at ACM-GECCO-2018 for the “Evolutionary Numerical Optimization” track, Kyoto, Japan

- Nikolaus Hansen: co-track chair at ACM-GECCO-2017 for the “Evolutionary Numerical Optimization” track, Berlin, Germany

9.1.2.2. Member of the Conference Program Committees or Reviewer

- Dimo Brockhoff reviewed for ACM-GECCO
- Anne Auger is reviewer for ACM-GECCO, ACM-FOGA , NIPS, ICML

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- Anne Auger and Nikolaus Hansen, members of the editorial board of the Evolutionary Computation Journal
- Dimo Brockhoff, co-guest editor of a special issue on Evolutionary Multiobjective Optimization in the Computers & Operations Research journal (issue 79), together with Joshua Knowles, Boris Naujoks, and Karthik Sindhya

9.1.3.2. Reviewer - Reviewing Activities

- Dimo Brockhoff reviewed in 2017 for IEEE Transactions on Evolutionary Computation, the Evolutionary Computation Journal, Natural Computing, PLoS One, Algorithmica, and Optimal Control, Applications and Methods
- Anne Auger reviewed in 2017 for IEEE Transactions on Evolutionary Computation, the Evolutionary Computation Journal, Algorithmica, SIAM Journal on Optimization

9.1.4. Invited Talks and Tutorials

- Dimo Brockhoff: invited tutorial on benchmarking (multiobjective) optimizers at the Symposium on Search-based Software Engineering (SSBSE’2017) in September 2017 in Paderborn, Germany
- Dimo Brockhoff: two invited talks (one on Evolutionary Multiobjective Optimization, one on Benchmarking) at the Mascot-Num conference in March 2017 in Paris, France
- Anne Auger and Nikolaus Hansen, tutorial on *Introduction to randomized continuous optimization* at the ACM-GECCO conference, Berlin, Germany
- Dimo Brockhoff: introductory tutorial on Evolutionary Multiobjective Optimization at the ACM-GECCO conference, Berlin, Germany
- Nikolaus Hansen: tutorial “A Practical Guide to Benchmarking and Experimentation” at the ACM-GECCO conference, Berlin, Germany
- Nikolaus Hansen: tutorial “CMA-ES and Advanced Adaptation Mechanisms” at the ACM-GECCO conference, Berlin, Germany, together with Youhei Akimoto

9.1.5. Leadership within the Scientific Community

- Since 2011, Anne Auger **Elected** member of the **ACM-SIGEVO executive board**, re-elected in 2017.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master: Dimo Brockhoff, “Introduction to Optimization”, 31.5h ETD, M2, Université Paris-Sud, France

Master: Anne Auger and Dimo Brockhoff, “Advanced Optimization”, 31.5h ETD, M2, Université Paris-Sud, France

Master: Anne Auger, “Derivative-free Optimization", Paris-Saclay, Optimization Master

Master: Anne Auger : Anne Auger (Introduction to Machine Learning, Advanced Machine Learning), Ecole Polytechnique, c.a. 50h

Summer school: Anne Auger and Dimo Brockhoff, July 3-7, 2017. CEA-EDF-Inria **summer school** on *Design and optimization under uncertainty of large-scale numerical models*. Course on *Introduction to Randomized Black-Box Numerical Optimization and CMA-ES*, Paris.

9.2.2. Supervision

PhD in progress: Cheikh Touré, topic: multiobjective optimization, started in October 2017, supervised by Anne Auger and Dimo Brockhoff

PhD in progress: Konstantinos Varelas, topic: constrained and expensive optimization, started in December 2017, supervised by Anne Auger and Dimo Brockhoff

9.2.3. Juries

- A. Auger in the PhD jury of Paul Feliot (defense July 2017)

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] O. AIT ELHARA.*Stochastic Black-Box Optimization and Benchmarking in Large Dimensions*, Université Paris-Saclay, July 2017, https://tel.archives-ouvertes.fr/tel-01615829.

Articles in International Peer-Reviewed Journal

[2] A. CHOTARD, A. AUGER.*Verifiable Conditions for the Irreducibility and Aperiodicity of Markov Chains by Analyzing Underlying Deterministic Models*, in "Bernoulli", 2017, https://arxiv.org/abs/1508.01644 , https://hal.inria.fr/hal-01222222.

[3] Y. OLLIVIER, L. ARNOLD, A. AUGER, N. HANSEN.*Information-Geometric Optimization Algorithms: A Unifying Picture via Invariance Principles*, in "Journal of Machine Learning Research", 2017, vol. 18, n^{o} 18, p. 1-65, https://hal.inria.fr/hal-01515898.

International Conferences with Proceedings

[4] Y. AKIMOTO, A. AUGER, N. HANSEN.*Quality Gain Analysis of the Weighted Recombination Evolution Strategy on General Convex Quadratic Functions*, in "Proceedings of the 14th ACM/SIGEVO Conference on Foundations of Genetic Algorithms, FOGA", Copenhagen, Denmark, January 2017, p. 111-126 [*DOI :* 10.1145/3040718.3040720], https://hal.inria.fr/hal-01516326.

[5] A. ATAMNA, A. AUGER, N. HANSEN.*Linearly Convergent Evolution Strategies via Augmented Lagrangian Constraint Handling*, in "The 14th ACM/SIGEVO Workshop on Foundations of Genetic Algorithms (FOGA XIV)", Copenhagen, Denmark, January 2017, p. 149 - 161 [*DOI :* 10.1145/3040718.3040732], https://hal.inria.fr/hal-01455379.

[6] D. BROCKHOFF, A. AUGER, N. HANSEN, T. TUŠAR.*Quantitative Performance Assessment of Multiobjective Optimizers: The Average Runtime Attainment Function*, in "Evolutionary Multi-Criterion Optimization (EMO 2017)", Münster, Germany, LNCS, March 2017, vol. 10173, p. 103-119 [*DOI :* 10.1007/978-3-319-54157-0_8], https://hal.inria.fr/hal-01591151.

[7] D. M. NGUYEN, N. HANSEN.*Benchmarking CMAES-APOP on the BBOB noiseless testbed*, in "Proceedings of the 2017 Genetic and Evolutionary Computation Conference Companion (GECCO '17 Companion)", Berlin, Germany, July 2017, p. 1756 - 1763 [*DOI :* 10.1145/2908812.2908864], https://hal.inria.fr/hal-01591423.

[8] T. TUŠAR, N. HANSEN, D. BROCKHOFF.*Anytime Benchmarking of Budget-Dependent Algorithms with the COCO Platform*, in "IS 2017 - International multiconference Information Society", Ljubljana, Slovenia, October 2017, p. 1-4, https://hal.inria.fr/hal-01629087.

Other Publications

[9] Y. AKIMOTO, A. AUGER, N. HANSEN.*Quality Gain Analysis of the Weighted Recombination Evolution Strategy on General Convex Quadratic Functions*, December 2017, Submitted to Journal of Theoretical Computer Science, https://hal.inria.fr/hal-01662568.

[10] A. ATAMNA, A. AUGER, N. HANSEN.*On Invariance and Linear Convergence of Evolution Strategies with Augmented Lagrangian Constraint Handling*, December 2017, working paper or preprint, https://hal.inria.fr/hal-01660728.

References in notes

[11] Y. AKIMOTO, N. HANSEN.*Online model selection for restricted covariance matrix adaptation*, in "International Conference on Parallel Problem Solving from Nature", Springer, 2016, p. 3–13.

[12] Y. AKIMOTO, N. HANSEN.*Projection-based restricted covariance matrix adaptation for high dimension*, in "Proceedings of the 2016 on Genetic and Evolutionary Computation Conference", ACM, 2016, p. 197–204.

[13] D. V. ARNOLD, J. PORTER.*Towards au Augmented Lagrangian Constraint Handling Approach for the* $(1+1)$*-ES*, in "Genetic and Evolutionary Computation Conference", ACM Press, 2015, p. 249-256.

[14] A. ATAMNA, A. AUGER, N. HANSEN.*Linearly Convergent Evolution Strategies via Augmented Lagrangian Constraint Handling*, in "Foundation of Genetic Algorithms (FOGA)", 2017.

[15] A. AUGER, N. HANSEN.*Linear Convergence of Comparison-based Step-size Adaptive Randomized Search via Stability of Markov Chains*, in "SIAM Journal on Optimization", 2016, vol. 26, n^{o} 3, p. 1589-1624.

[16] V. S. BORKAR.*Stochastic approximation: a dynamical systems viewpoint*, 2008, Cambridge University Press.

[17] V. BORKAR, S. MEYN.*The O.D.E. Method for Convergence of Stochastic Approximation and Reinforcement Learning*, in "SIAM Journal on Control and Optimization", January 2000, vol. 38, n^{o} 2.

[18] C. A. COELLO COELLO.*Constraint-handling techniques used with evolutionary algorithms*, in "Proceedings of the 2008 Genetic and Evolutionary Computation Conference", ACM, 2008, p. 2445–2466.

[19] G. COLLANGE, S. REYNAUD, N. HANSEN.*Covariance Matrix Adaptation Evolution Strategy for Multidisciplinary Optimization of Expendable Launcher Families*, in "13th AIAA/ISSMO Multidisciplinary Analysis Optimization Conference, Proceedings", 2010.

[20] J. E. DENNIS, R. B. SCHNABEL.*Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1983.

[21] N. HANSEN, A. OSTERMEIER.*Completely Derandomized Self-Adaptation in Evolution Strategies*, in "Evolutionary Computation", 2001, vol. 9, n^{o} 2, p. 159–195.

[22] J. N. HOOKER.*Testing heuristics: We have it all wrong*, in "Journal of heuristics", 1995, vol. 1, n^{o} 1, p. 33–42.

[23] I. KRIEST, V. SAUERLAND, S. KHATIWALA, A. SRIVASTAV, A. OSCHLIES.*Calibrating a global three-dimensional biogeochemical ocean model (MOPS-1.0)*, in "Geoscientific Model Development", 2017, vol. 10, n^{o} 1, 127.

[24] H. J. KUSHNER, G. YIN.*Stochastic approximation and recursive algorithms and applications*, Applications of mathematics, Springer, New York, 2003, http://opac.inria.fr/record=b1099801.

[25] P. MACALPINE, S. BARRETT, D. URIELI, V. VU, P. STONE.*Design and Optimization of an Omnidirectional Humanoid Walk: A Winning Approach at the RoboCup 2011 3D Simulation Competition*, in "Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence (AAAI)", July 2012.

[26] S. MEYN, R. TWEEDIE.*Markov Chains and Stochastic Stability*, Springer-Verlag, New York, 1993.

[27] T. SALIMANS, J. HO, X. CHEN, I. SUTSKEVER.*Evolution strategies as a scalable alternative to reinforcement learning*, in "arXiv preprint arXiv:1703.03864", 2017.

[28] J. UHLENDORF, A. MIERMONT, T. DELAVEAU, G. CHARVIN, F. FAGES, S. BOTTANI, G. BATT, P. HERSEN.*Long-term model predictive control of gene expression at the population and single-cell levels*, in "Proceedings of the National Academy of Sciences", 2012, vol. 109, n^{o} 35, p. 14271–14276.

Project-Team SELECT

Model selection in statistical learning

IN COLLABORATION WITH: Laboratoire de mathématiques d'Orsay de l'Université de Paris-Sud (LMO)

IN PARTNERSHIP WITH:
CNRS

Université Paris-Sud (Paris 11)

RESEARCH CENTER
Saclay - Île-de-France

THEME
Optimization, machine learning and statistical methods

Table of contents

Project-Team SELECT

Creation of the Project-Team: 2007 January 01

Keywords:

Computer Science and Digital Science:

A3.1.1. - Modeling, representation
A3.1.8. - Big data (production, storage, transfer)
A3.2.2. - Knowledge extraction, cleaning
A3.3.2. - Data mining
A3.3.3. - Big data analysis
A3.4.1. - Supervised learning
A3.4.2. - Unsupervised learning
A3.4.3. - Reinforcement learning
A3.4.4. - Optimization and learning
A3.4.5. - Bayesian methods
A3.4.6. - Neural networks
A3.4.7. - Kernel methods
A3.4.8. - Deep learning
A5.3.3. - Pattern recognition
A6.2.4. - Statistical methods
A6.2.6. - Optimization

Other Research Topics and Application Domains:

B1.1.5. - Genetics
B1.1.6. - Genomics
B1.1.9. - Bioinformatics
B1.1.10. - Mathematical biology
B9.4.2. - Mathematics

1. Personnel

Research Scientists

Sylvain Arlot [CNRS, Researcher]
Benjamin Auder [CNRS, Researcher]
Kevin Bleakley [Inria, Researcher]
Gilles Celeux [Inria, Emeritus, HDR]
Matthieu Lerasle [CNRS, Researcher]

Faculty Members

Pascal Massart [Team leader, Univ Paris-Sud, Professor, HDR]
Christine Keribin [Univ Paris-Sud, Associate Professor]
Claire Lacour [Univ Paris-Sud, Associate Professor]
Patrick Pamphile [Univ Paris-Sud, Associate Professor]
Jean-Michel Poggi [Univ René Descartes Paris, Professor, HDR]

Technical Staff

Josselin Demont [Inria, until Jan 2017]

Jonas Renault [Inria, until Sep 2017]
Christian Poli [Inria]

PhD Students
Benjamin Goehry [Univ Paris-Sud]
Valérie Robert [Univ Paris-Sud, until Aug 2017]
Yann Vasseur [Univ Paris-Sud, until Aug 2017]
Neska El Haouij [Univ Paris-Sud]
Hedi Hadiji [Univ Paris-Sud]
Minh Lien Nguyen [Univ Paris-Sud]
Florence Ducros [Univ Paris-Sud]

Post-Doctoral Fellow
Kaniav Kamary [Inria, until Jul 2017]

Administrative Assistant
Olga Mwana Mobulakani [Inria]

2. Overall Objectives

2.1. Model selection in Statistics

The research domain for the SELECT project is statistics. Statistical methodology has made great progress over the past few decades, with a variety of statistical learning software packages that support many different methods and algorithms. Users now face the problem of choosing among them, to select the most appropriate method for their data sets and objectives. The problem of model selection is an important but difficult problem, both theoretically and practically. Classical model selection criteria, which use penalized minimum-contrast criteria with fixed penalties, are often based on unrealistic assumptions.

SELECT aims to provide efficient model selection criteria with data-driven penalty terms. In this context, SELECT aims to improve the toolkit of statistical model selection criteria from both theoretical and practical perspectives. Currently, SELECT is focusing its effort on variable selection in statistical learning, hidden-structure models and supervised classification. Its domains of application concern reliability, curve classification, phylogenetic analysis and classification in genetics. New developments in SELECT activities are concerned with applications in biostatistics (statistical analysis of medical images) and biology.

3. Research Program

3.1. General presentation

From applications we treat on a day-to-day basis, we have learned that some assumptions currently used in asymptotic theory for model selection are often irrelevant in practice. For instance, it is not realistic to assume that the target belongs to the family of models in competition. Moreover, in many situations, it is useful to make the size of the model depend on the sample size, which makes asymptotic analyses breakdown. An important aim of SELECT is to propose model selection criteria which take such practical constraints into account.

3.2. A nonasymptotic view of model selection

An important goal of SELECT is to build and analyze penalized log-likelihood model selection criteria that are efficient when the number of models in competition grows to infinity with the number of observations. Concentration inequalities are a key tool for this, and lead to data-driven penalty choice strategies. A major research direction for SELECT consists of deepening the analysis of data-driven penalties, both from the theoretical and practical points of view. There is no universal way of calibrating penalties, but there are several different general ideas that we aim to develop, including heuristics derived from Gaussian theory, special strategies for variable selection, and resampling methods.

3.3. Taking into account the modeling purpose in model selection

Choosing a model is not only difficult theoretically. From a practical point of view, it is important to design model selection criteria that accommodate situations in which the data probability distribution P is unknown, and which take the model user's purpose into account. Most standard model selection criteria assume that P belongs to one of a set of models, without considering the purpose of the model. By also considering the model user's purpose, we can avoid or overcome certain theoretical difficulties, and produce flexible model selection criteria with data-driven penalties. The latter is useful in supervised classification and hidden-structure models.

3.4. Bayesian model selection

The Bayesian approach to statistical problems is fundamentally probabilistic: a joint probability distribution is used to describe the relationships among all unknowns and the data. Inference is then based on the posterior distribution, i.e., the conditional probability distribution of the parameters given the observed data. Exploiting the internal consistency of the probability framework, the posterior distribution extracts relevant information in the data and provides a complete and coherent summary of post-data uncertainty. Using the posterior to solve specific inference and decision problems is then straightforward, at least in principle.

4. Application Domains

4.1. Introduction

A key goal of SELECT is to produce methodological contributions in statistics. For this reason, the SELECT team works with applications that serve as an important source of interesting practical problems and require innovative methodology to address them. Many of our applications involve contracts with industrial partners, e.g., in reliability, although we also have several academic collaborations, e.g., in genetics and image analysis.

4.2. Curve classification

The field of classification for complex data such as curves, functions, spectra and time series, is an important problem in current research. Standard data analysis questions are being looked into anew, in order to define novel strategies that take the functional nature of such data into account. Functional data analysis addresses a variety of applied problems, including longitudinal studies, analysis of fMRI data, and spectral calibration.

We are focused in particular on unsupervised classification. In addition to standard questions such as the choice of the number of clusters, the norm for measuring the distance between two observations, and vectors for representing clusters, we must also address a major computational problem: the functional nature of the data, which requires new approaches.

4.3. Computer experiments and reliability

For several years now, SELECT has collaborated with the EDF-DER *Maintenance des Risques Industriels* group. One important theme involves the resolution of inverse problems using simulation tools to analyze incertainty in highly complex physical systems.

The other major theme concerns reliability, through a research collaboration with Nexter involving a Cifre convention. This collaboration concerns a lifetime analysis of a vehicle fleet to assess ageing.

Moreover, a collaboration is ongoing with Dassault Aviation on the modal analysis of mechanical structures, which aims to identify the vibration behavior of structures under dynamic excitation. From the algorithmic point of view, modal analysis amounts to estimation in parametric models on the basis of measured excitations and structural response data. In literature and existing implementations, the model selection problem associated with this estimation is currently treated by a rather weighty and heuristic procedure. In the context of our own research, model selection via penalization methods are being tested on this model selection problem.

4.4. Analysis of genomic data

For many years now, SELECT collaborates with Marie-Laure Martin-Magniette (URGV) for the analysis of genomic data. An important theme of this collaboration is using statistically sound model-based clustering methods to discover groups of co-expressed genes from microarray and high-throughput sequencing data. In particular, identifying biological entities that share similar profiles across several treatment conditions, such as co-expressed genes, may help identify groups of genes that are involved in the same biological processes.

Yann Vasseur has completed a thesis co-supervised by Gilles Celeux and Marie-Laure Martin-Magniette on this topic, which is also an interesting investigation domain for the latent block model developed by SELECT. For this work, Yann Vasseur dealt with high-dimensional ill-posed problems where the number of variable was almost equal to the number of observations. He designed heuristic tools using regularized regression methods to circumvent this difficulty.

SELECT collaborates with Anavaj Sakuntabhai and Philippe Dussart (Pasteur Institute) on predicting dengue severity using only low-dimensional clinical data obtained at hospital arrival. An ongoing project also involves statistical meta-analysis of newly collected dengue gene expression data along with recently published data sets from other groups. Further collaborations are underway in dengue fever and encephalitis with researchers at the Pasteur Institute.

SELECT collaborates with Inserm/Paris-Saclay researchers at Kremlin-Bicêtre hospital on cyclic transcriptional clocks and renal corticosteroid signaling, developing statistical tests for synchronous signals.

SELECT is involved in the ANR “jeunes chercheurs” MixStatSeq directed by Cathy Maugis (INSA Toulouse), which is concerned with statistical analysis and clustering of RNASeq genomics data.

4.5. Pharmacovigilance

A collaboration is ongoing with Pascale Tubert-Bitter, Ismael Ahmed and Mohamed Sedki (Pharmacoepidemiology and Infectious Diseases, PhEMI) for the analysis of pharmacovigilance data. In this framework, the goal is to detect, as soon as possible, potential associations between certain drugs and adverse effects, which appeared after the authorized marketing of these drugs. Instead of working on aggregate data (contingency table) like is usually the case, the approach developed aims to deal with individual’s data, which perhaps gives more information. Valerie Robert has completed a thesis co-supervised by Gilles Celeux and Christine Keribin on this topic, which involved the development of a new model-based clustering method, inspired by latent block models. Morever, she has defined new tools to estimate and assess the block clustering involved in these models.

4.6. Spectroscopic imaging analysis of ancient materials

Ancient materials, encountered in archaeology and paleontology are often complex, heterogeneous and poorly characterized before physico-chemical analysis. A popular technique to gather as much physico-chemical information as possible, is spectro-microscopy or spectral imaging, where a full spectra, made of more than a thousand samples, is measured for each pixel. The produced data is tensorial with two or three spatial dimensions and one or more spectral dimensions, and requires the combination of an “image” approach with a “curve analysis” approach. Since 2010 SELECT, collaborates with Serge Cohen (IPANEMA) on the development of conditional density estimation through GMM, and non-asymptotic model selection, to perform stochastic segmentation of such tensorial datasets. This technique enables the simultaneous accounting for spatial and spectral information, while producing statistically sound information on morphological and physico-chemical aspects of the studied samples.

5. New Software and Platforms

5.1. BlockCluster

Block Clustering

KEYWORDS: Statistic analysis - Clustering package
SCIENTIFIC DESCRIPTION: Simultaneous clustering of rows and columns, usually designated by biclustering, co-clustering or block clustering, is an important technique in two way data analysis. It consists of estimating a mixture model which takes into account the block clustering problem on both the individual and variables sets. The blockcluster package provides a bridge between the C++ core library and the R statistical computing environment. This package allows to co-cluster binary, contingency, continuous and categorical data-sets. It also provides utility functions to visualize the results. This package may be useful for various applications in fields of Data mining, Information retrieval, Biology, computer vision and many more.
FUNCTIONAL DESCRIPTION: BlockCluster is an R package for co-clustering of binary, contingency and continuous data based on mixture models.

- Participants: Christophe Biernacki, Gilles Celeux, Parmeet Bhatia, Serge Iovleff, Vincent Brault and Vincent Kubicki
- Partner: Université de Technologie de Compiègne
- Contact: Serge Iovleff
- URL: http://cran.r-project.org/web/packages/blockcluster/index.html

5.2. MASSICCC

Massive Clustering with Cloud Computing
KEYWORDS: Statistic analysis - Big data - Machine learning - Web Application
SCIENTIFIC DESCRIPTION: The web application let users use several software packages developed by Inria directly in a web browser. Mixmod is a classification library for continuous and categorical data. MixtComp allows for missing data and a larger choice of data types. BlockCluster is a library for co-clustering of data. When using the web application, the user can first upload a data set, then configure a job using one of the libraries mentioned and start the execution of the job on a cluster. The results are then displayed directly in the browser allowing for rapid understanding and interactive visualisation.
FUNCTIONAL DESCRIPTION: The MASSICCC web application offers a simple and dynamic interface for analysing heterogeneous data with a web browser. Various software packages for statistical analysis are available (Mixmod, MixtComp, BlockCluster) which allow for supervised and supervised classification of large data sets.

- Contact: Jonas Renault
- URL: https://massiccc.lille.inria.fr

5.3. Mixmod

Many-purpose software for data mining and statistical learning
KEYWORDS: Data modeling - Mixed data - Classification - Data mining - Big data
FUNCTIONAL DESCRIPTION: Mixmod is a free toolbox for data mining and statistical learning designed for large and highdimensional data sets. Mixmod provides reliable estimation algorithms and relevant model selection criteria.

It has been successfully applied to marketing, credit scoring, epidemiology, genomics and reliability among other domains. Its particularity is to propose a model-based approach leading to a lot of methods for classification and clustering.

Mixmod allows to assess the stability of the results with simple and thorough scores. It provides an easy-to-use graphical user interface (mixmodGUI) and functions for the R (Rmixmod) and Matlab (mixmodForMatlab) environments.

- Participants: Benjamin Auder, Christophe Biernacki, Florent Langrognet, Gérard Govaert, Gilles Celeux, Remi Lebret and Serge Iovleff
- Partners: CNRS - Université Lille 1 - LIFL - Laboratoire Paul Painlevé - HEUDIASYC - LMB
- Contact: Gilles Celeux
- URL: http://www.mixmod.org

6. New Results

6.1. Model selection in Regression and Classification

Participants: Gilles Celeux, Pascal Massart, Sylvain Arlot, Jean-Michel Poggi, Kevin Bleakley.

The well-documented and consistent variable selection procedure in model-based cluster analysis and classification that Cathy Maugis (INSA Toulouse) designed during her PhD thesis in SELECT, makes use of stepwise algorithms which are painfully slow in high dimensions. In order to circumvent this drawback, Gilles Celeux, in collaboration with Mohammed Sedki (Université Paris XI) and Cathy Maugis, have recently submitted an article where variables are sorted using a lasso-like penalization adapted to the Gaussian mixture model context. Using this ranking to select variables, they avoid the combinatory problem of stepwise procedures. The performances on challenging simulated and real data sets are similar to the standard procedure, with a CPU time divided by a factor of more than a hundred.

In collaboration with Jean-Michel Marin (Université de Montpellier) and Olivier Gascuel (LIRMM), Gilles Celeux has continued research aiming to select a short list of models rather a single model. This short list is declared to be compatible with the data using a p-value derived from the Kullback-Leibler distance between the model and the empirical distribution. Furthermore, the Kullback-Leibler distances at hand are estimated through nonparametric and parametric bootstrap procedures. Different strategies are compared through numerical experiments on simulated and real data sets.

6.2. Estimator selection and statistical tests

Participants: Sylvain Arlot, Matthieu Lerasle.

G. Maillard, S. Arlot and M. Lerasle studied a method mixing cross-validation with aggregation, called aggregated hold-out (Agghoo), which is already used by several practitioners. Agghoo can also be related to bagging. According to numerical experiments, Agghoo can improve significantly cross-validation's prediction error, at the same computational cost; this makes it very promising as a general-purpose tool for prediction. This work provides the first theoretical guarantees on Agghoo, in the supervised classification setting, ensuring that one can use it safely: at worst, Agghoo performs like hold-out, up to a constant factor. A non-asymptotic oracle inequality is also proved, in binary classification under the margin condition, which is sharp enough to get (fast) minimax rates.

With G. Lecué, Matthieu Lerasle working on "learning from MOM's principles", showing that a recent procedure by Lugosi and Mendelson can be derived by applying Le Cam's "estimation from tests" procedure to MOM's tests. They also showed some robustness properties of these estimators, proving that the rates of convergence of this estimator are not downgraded even if some "outliers" have corrupted the dataset, and the other data have only first and second moments equal to that of the targeted probability distribution.

6.3. Statistical learning methodology and theory

Participants: Gilles Celeux, Serge Cohen, Christine Keribin, Michel Prenat, Kaniav Kamary, Sylvain Arlot, Benjamin Auder, Jean-Michel Poggi, Neska El Haouij, Kevin Bleakley, Matthieu Lerasle.

Gilles Celeux and Serge Cohen have started research in collaboration with Agnès Grimaud (UVSQ) to perform clustering of hyperspectral images which respects spatial constraints. This is a one-class classification problem where distances between spectral images are given by the χ^2 distance, while spatial homogeneity is associated with a single link distance.

Gilles Celeux continued his collaboration with Jean-Patrick Baudry on model-based clustering. This year, they started work on assessing model-based clustering methods on cytometry data sets. The interest of these is that they involve combining clustering and classification tasks in a unified framework.

Gillies Celeux and Julie Josse have started research on missing data for model-based clustering in collaboration with Christophe Biernacki (Modal, Inria Lille). This year, they have proposed a model for mixture analysis involving not missing-at-random mixtures.

In the framework of MASSICCC, Benjamin Auder and Gilles Celeux have started research on the graphical representation of model-based clusters. The aim of this is to better-display proximity between clusters.

For a long time unsolved, the consistency and asymptotic normality of the maximum likelihood and variational estimators of the latent block model were finally tackled and obtained in a joint work with V. Brault and M. Mariadassou.

J-M. Poggi (with R. Genuer, C. Tuleau-Malot, N. Villa-Vialaneix), have published an article on random forests in “big data” classification problems, and have performed a review of available proposals about random forests in parallel environments as well as on online random forests. Three variants involving subsampling, Big Data-bootstrap and MapReduce respectively were tested on two massive datasets, one simulated one, and the other, real-world data.

With G. Lecué, Matthieu Lerasle worked on robust machine learning by median-of-means, providing an alternative to the Lugosi and Mendelson approach based on median of means for learning. This alternative is easier to present and to analyse theoretically. Furthermore, they proposed an algorithm to approximate this estimator, which could not be done for Lugosi and Mendelson’s champions of tournaments (submitted).

6.4. Estimation for conditional densities in high dimension

Participants: Claire Lacour, Jeanne Nguyen.

Jeanne Nguyen is working on estimation for conditional densities in high dimension. Much more informative than the regression function, conditional densities are of high interest in recent methods, particularly in the Bayesian framework (studying the posterior distribution). Considering a specific family of kernel estimators, she is studying a greedy algorithm for selecting the bandwidth. Her method addresses several issues: avoiding the curse of high dimensionality under some suitably defined sparsity conditions, being computationally efficient using iterative procedures, and early variable selection, providing theoretical guarantees on the minimax risk.

6.5. Reliability

Participants: Gilles Celeux, Florence Ducros, Patrick Pamphile.

Since June 2015, in the framework of a CIFRE convention with Nexter, Florence Ducros has begun a thesis on the modeling of aging of vehicles, supervised by Gilles Celeux and Patrick Pamphile. This thesis should lead to designing an efficient maintenance strategy according to vehicle use profiles. Moreover, warranty cost calculations are made in the context of heterogeneous usages. This required estimations of mixtures and competing risk models in a highly-censored setting.

This year, Patrick Pamphile and Florence Ducros have published an article which proposes a two-component Weibull mixture model for modelling unobserved heterogeneity in heavily censored lifetime data collection. Performance of classical estimation methods (maximum of likelihood, EM, full Bayes and MCMC) are poor due to the high number of parameters and the heavy censoring. Thus, a Bayesian bootstrap method called Bayesian Restoration Maximization, was used. Sampling from the posterior distribution was obtained thanks to an importance sampling technique. Simulation results showed that, even with heavy censoring, BRM is effective both in term of estimate’s precision and computation times.

6.6. Statistical analysis of genomic data

Participants: Gilles Celeux, Christine Keribin, Yann Vasseur, Kevin Bleakley.

The subject of Yann Vasseur’s PhD Thesis, supervised by Gilles Celeux and Marie-Laure Martin-Magniette (INRA URGV), was the inference of a regulatory network for Transcriptions Factors (TFs), which are specific genes, of *Arabidopsis thaliana*. For this, a transcriptome dataset with a similar number of TFs and statistical units was available. They reduced the dimension of the network to avoid high-dimensional difficulties. Representing this network with a Gaussian graphical model, the following procedure was defined:

1. *Selection step*: choose the set of TF regulators (supports) of each TF.
2. *Classification step*: deduce co-factor groups (TFs with similar expression levels) from these supports.

Thus, the reduced network would be built on the co-factor groups. Currently, several selection methods based on Gauss-LASSO and resampling procedures have been applied to the dataset. The study of stability and parameter calibration of these methods is in progress. The TFs are clustered with the Latent Block Model into a number of co-factor groups, selected with BIC or the exact ICL criterion. Since these models are built in an ad hoc way, Yann Vasseur has defined complex simulation tools to asses their performances in a proper way.

In collaboration with Benno Schwikowski, Iryna Nikolayeva and Anavaj Sakuntabhai (Pasteur Institute, Paris), Kevin Bleakley worked on using 2-d isotonic regression to predict dengue fever severity at hospital arrival using high-dimensional microarray gene expression data. Important marker genes for dengue severity have been detected, some of which now have been validated in external lab trials, and an article has now been submitted.

In collaboration with researchers from the Pasteur Institute, Kevin Bleakley worked on statistical tests in the context of research into what leads to dengue fever *without symptoms* as opposed to *with* symptoms. This work was published in *Science Translational Medicine*.

Kevin Bleakley has also collaborated with Inserm/Paris-Saclay researchers at Kremlin-Bicêtre hospital on cyclic transcriptional clocks and renal corticosteroid signaling, and has developed novel statistical tests for detecting synchronous signals. This work is submitted.

6.7. Model-based clustering for pharmacovigilance data

Participants: Gilles Celeux, Christine Keribin, Valérie Robert.

In collaboration with Pascale Tubert-Bitter, Ismael Ahmed and Mohamed Sedki, Gilles Celeux and Christine Keribin worked on the detection of associations between drugs and adverse events in the framework of the PhD of Valerie Robert, which was defended this year. At first, this team developed model-based clustering inspired by latent block models (LBMs), which consists of co-clustering rows and columns of two binary tables, imposing the same row ranking. This enabled it to highlight subgroups of individuals sharing the same drug profile, and subgroups of adverse effects and drugs with strong interactions. Furthermore, some sufficient conditions are provided to obtain identifiability of the model, and some results are shown for simulated data. The exact ICL criterion has been extended to this double block latent model. Through computer experiments, Valérie Robert demonstrated the interest of the proposed model, compared with standard contingency table analysis, to detect co-prescription and masking effects.

Futhermore, with V. Robert, C. Kerebin and G. Celeux showed that it can be useful to use an LBM model on a contingency table of drugs and adverse effects to do cluster initialization for dealing with individual's data.

6.8. Statistical rating and ranking of scientific journals

Participants: Gilles Celeux, Julie Josse, Jean-Louis Foulley.

In collaboration with Jean-Louis Foulley (Montpellier University), Gilles Celeux and Julie Josse have done research on the statistical rating and ranking of scientific journals. They have proposed Dirichlet multinomial Bayesian models for pagerank-type algorithms allowing self-citations to be excluded. The resulting methods were tested on a set of 47 scientific journals.

6.9. Statistical mathematics

Participant: Matthieu Lerasle.

In collaboration with R. Diel, Matthieu Lerasle published an article on nonparametric estimation for random walks in random environments. They proposed a non-parametric approach for estimating the distribution of the environment from the observation of one trajectory of a random walk in it. They obtained risk bounds in sup-norm for the cumulative distribution function of the environment.

6.10. Random graph theory

Participant: Matthieu Lerasle.

In collaboration with R. Chetrite and R. Diel, Matthieu Lerasle published an article on the number of potential winners in the Bradley-Terry model in random environments. They proposed the first mathematical study of the Bradley-Terry model where the values of players are i.i.d. realisations of some distribution. They proved that a Bradley-Terry tournament is fair (in the sense that the best player ends up with the largest number of victories) under a certain convexity condition on the tail distribution of the values. They also showed that this condition is sharp and provided sharp estimate of the number of potential winners when the condition fails.

He also collaborated with R. Diel and S. Le Corff on learning latent structures of large random graphs, investigating the possibility of estimating latent structure in sparsely observed random graphs. The main example was a Bradley-Terry tournament where each team has only played a few games. It is well known that individual values of the teams cannot be consistently estimated in this setting. They showed that their distribution on the other hand can be, and provide general tools for bounding the risk of the maximum likelihood estimator (submitted).

7. Bilateral Contracts and Grants with Industry

7.1. Contract with NEXTER

Participants: Gilles Celeux, Florence Ducros, Patrick Pamphile.

SELECT has a contract with Nexter regarding modeling the reliability of vehicles.

7.2. Bilateral Grants with Industry

Benjamin Auder and Jean-Michel Poggi are participants in the grant PGMO-IRSDI, in the *Research Initiative In Industrial Data Science* context, on the subject: Disaggregated Electricity Forecasting using Clustering of Individual Consumers.

8. Partnerships and Cooperations

8.1. Regional Initiatives

Gilles Celeux and Christine Keribin have a collaboration with the Pharmacoepidemiology and Infectious Diseases (PhEMI, INSERM) groups.

Sylvain Arlot and Pascal Massart co-organize a working group at ENS (Ulm) on statistical learning.

8.2. National Initiatives

8.2.1. ANR

SELECT is part of the ANR funded MixStatSeq.

8.3. International Initiatives

Gilles Celeux is one of the co-organizers of the international working group on model-based clustering. This year this workshop took place in Perugia, Italy

8.4. International Research Visitors

8.4.1. Visits to International Teams

8.4.1.1. Research Stays Abroad

Kevin Bleakley stayed at the Pasteur Institute, Cambodia, while working on several collaborations in dengue fever research, from late 2016 until early 2017.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

Sylvain Arlot organized (with Guillaume Charpiat) the Workshop Statistics/Learning at Paris-Saclay (2nd edition), at IHES (Bures-sur-Yvette).

9.1.1.2. Member of the Organizing Committees

- Gilles Celeux is one of the co-organizers of the international working group on model-based clustering. This year the workshop took place in Perugia, Italy.
- Sylvain Arlot is one of the co-organizers of the Junior Conference on Data Science and Engineering at Paris-Saclay (2nd edition in 2017).
- Jean-Michel Poggi was president of the Scientific Program Committee, ENBIS 2017, Naples, 10-14 June 2017.
- Jean-Michel Poggi was member of the Conference Scientific Board of IES 2017, Naples, Italy, 6-8 September 2017.

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards

Gilles Celeux is Editor-in-Chief of the *Journal de la SFdS*. He is Associate Editor of *Statistics and Computing*, *CSBIGS*.

Pascal Massart is Associate Editor of *Annals of Statistics*, *Confluentes Mathematici*, and *Foundations and Trends in Machine Learning*.

Jean-Michel Poggi is Associate Editor of *Journal of Statistical Software*, *Journal de la SFdS* and *CSBIGS*.

9.1.2.2. Reviewer - Reviewing Activities

The members of the team have reviewed numerous papers for numerous international journals.

9.1.3. Invited Talks

The members of the team have given many invited talks on their research in the course of 2016.

9.1.4. Leadership within the Scientific Community

Jean-Michel Poggi is:

- Vice-President ENBIS (European Network for Business and Industrial Statistics), 2015-18
- Vice-President FENStatS (Federation of European National Statistical Societies) since 2012
- Council Member of the ISI (2015-19)
- Member of the Board of Directors of the ERS of IASC (since 2014)

9.1.5. Scientific Expertise

Jean-Michel Poggi is member of the EMS Committee for Applied Mathematics (since 2014).

9.1.6. Research Administration

Jean-Michel Poggi is the president of ECAS (European Courses in Advanced Statistics) since 2015.

Sylvain Arlot coordinates (jointly with Marc Schoenauer, Inria Saclay) the math-STIC program of the Labex Mathématique Hadamard.

Christine Keribin is treasurer of the Société Française de Statistique (SFdS).

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

SELECT members teach various courses at several different universities, and in particular the Master 2 "Mathématique de l'aléatoire" of Université Paris-Saclay.

9.2.2. Supervision

PhD: Valérie Robert, 2013, Gilles Celeux and Christine Keribin. Defended in June 2017

PhD : Yann Vasseur, 2013, Gilles Celeux and Marie-Laure Martin-Magniette (URGV). Defended in December 2017

PhD in progress: Neska El Haouij, 2014, Jean-Michel Poggi and Meriem Jaïdane, Raja Ghozi (ENIT Tunisie) and Sylvie Sevestre-Ghalila (CEA LinkLab), Thesis ENITUPS

PhD in progress: Florence Ducros, 2015, Gilles Celeux and Patrick Pamphile

PhD in progress: Claire Brécheteau, 2015, Pascal Massart

PhD in progress: Hedi Hadiji, 2017, Pascal Massart

PhD in progress: Eddie Aamari, 2015, Pascal Massart and Frédéric Chazal

PhD: Damien Garreau, 2013, Sylvain Arlot and Gérard Biau (UPMPC). Defended in October 2017

PhD in progress: Guillaume Maillard, 2016, Sylvain Arlot and Matthieu Lerasle

PhD in progress: Jeanne Nguyen, 2015, Claire Lacour and Vincent Rivoirard (Univ Paris Dauphine)

PhD in progress: Benjamin Goehry, 2015, Pascal Massart and Jean-Michel Poggi

Masters internship: Thomas Prochwicz. Christine Keribin conducted a preliminary study on expert aggregation by supervising this three month internship.

9.2.3. Juries

S. Arlot was a member of the Ph.D. jury of Jilai Mei (Université Paris-Sud).

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] E. AAMARI. *Convergence Rates for Geometric Inference*, Université Paris-Saclay, September 2017, https://hal.inria.fr/tel-01607782.

Articles in International Peer-Reviewed Journal

[2] C. LACOUR, P. MASSART, V. RIVOIRARD.*Estimator selection: a new method with applications to kernel density estimation*, in "Sankhya A", August 2017, vol. 79, n^{o} 2, p. 298 - 335, https://arxiv.org/abs/1607.05091 [*DOI :* 10.1007/s13171-017-0107-5], https://hal.archives-ouvertes.fr/hal-01346081.

[3] E. SIMON-LORIERE, V. DUONG, A. TAWFIK, S. UNG, S. LY, I. CASADEMONT, M. PROT, N. COURTEJOIE, K. BLEAKLEY, P. BUCHY, A. TARANTOLA, P. DUSSART, T. CANTAERT, A. SAKUNTABHAI.*Increased adaptive immune responses and proper feedback regulation protect against clinical dengue*, in "Science Translational Medicine", August 2017, vol. 9, n^{o} 405, eaal5088, https://arxiv.org/abs/1712.05692 [*DOI :* 10.1126/SCITRANSLMED.AAL5088], https://hal.inria.fr/hal-01656594.

Invited Conferences

[4] J.-P. BAUDRY, G. CELEUX.*Assessing model-based clustering methods with cytometry data sets*, in "IFCS 2017 - Conference of the International Federation of Classification Societies", Tokyo, Japan, August 2017, https://hal.inria.fr/hal-01649085.

International Conferences with Proceedings

[5] C. KERIBIN, G. CELEUX, V. ROBERT.*The Latent Block Model: a useful model for high dimensional data*, in "ISI 2017 - 61st world statistics congress", Marrakech, Morocco, July 2017, p. 1-6, https://hal.inria.fr/hal-01658589.

Conferences without Proceedings

[6] V. BRAULT, A. CHANNAROND, V. ROBERT.*Généralisation de l'algorithme Largest Gaps pour le modèle des blocs latents non-paramétrique*, in "49èmes Journées de Statistique", Avignon, France, May 2017, https://hal.archives-ouvertes.fr/hal-01510984.

[7] V. BRAULT, C. KERIBIN, M. MARIADASSOU.*Équivalence asymptotique des vraisemblances observée et complète dans le modèle de blocs latents*, in "XXIV èmes Rencontres de la Société Francophone de Classification", Lyon, France, Société Francophone de Classification, June 2017, https://hal.archives-ouvertes.fr/hal-01510994.

Other Publications

[8] E. AAMARI, J. KIM, F. CHAZAL, B. MICHEL, A. RINALDO, L. WASSERMAN.*Estimating the Reach of a Manifold*, May 2017, https://arxiv.org/abs/1705.04565 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01521955.

[9] E. AAMARI, C. LEVRARD.*Non-Asymptotic Rates for Manifold, Tangent Space, and Curvature Estimation*, April 2017, https://arxiv.org/abs/1705.00989 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01516032.

[10] E. AAMARI, C. LEVRARD.*Stability and Minimax Optimality of Tangential Delaunay Complexes for Manifold Reconstruction*, November 2017, https://arxiv.org/abs/1512.02857 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01245479.

[11] S. ARLOT.*Cross-validation*, March 2017, https://arxiv.org/abs/1703.03167 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01485508.

[12] S. ARLOT.*Tutorial on statistical learning*, March 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01485506.

[13] V. BRAULT, C. KERIBIN, M. MARIADASSOU.*Consistency and Asymptotic Normality of Latent Blocks Model Estimators*, April 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01511960.

[14] N. EL HAOUIJ, J.-M. POGGI, R. E. GHOZI, S. SEVESTRE-GHALILA, M. JAÏDANE.*Random Forest-Based Approach for Physiological Functional Variable Selection: Towards Driver's Stress Level Classification*, January 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01426752.

[15] J.-L. FOULLEY, G. CELEUX, J. JOSSE.*Empirical Bayes approaches to PageRank type algorithms for rating scientific journals*, June 2017, working paper or preprint, https://hal.inria.fr/hal-01535134.

[16] G. MAILLARD, S. ARLOT, M. LERASLE.*Cross-validation improved by aggregation: Agghoo*, September 2017, https://arxiv.org/abs/1709.03702 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01585595.

[17] V. ROBERT, Y. VASSEUR.*Comparing high dimensional partitions with the Coclustering Adjusted Rand Index*, May 2017, working paper or preprint, https://hal.inria.fr/hal-01524832.

Project-Team SPECFUN

Symbolic Special Functions : Fast and Certified

RESEARCH CENTER
Saclay - Île-de-France

THEME
Algorithmics, Computer Algebra and Cryptology

Table of contents

Project-Team SPECFUN

Creation of the Team: 2012 November 01, updated into Project-Team: 2014 July 01

Keywords:

Computer Science and Digital Science:

A2.1.10. - Domain-specific languages
A2.1.11. - Proof languages
A2.4.3. - Proofs
A4.5. - Formal methods for security
A7.2. - Logic in Computer Science
A8.1. - Discrete mathematics, combinatorics
A8.4. - Computer Algebra
A8.5. - Number theory
A8.9. - Performance evaluation

Other Research Topics and Application Domains:

B9.4.2. - Mathematics
B9.4.3. - Physics

1. Personnel

Research Scientists

Frédéric Chyzak [Team leader, Inria, Researcher, HDR]
Alin Bostan [Inria, Researcher, HDR]
Philippe Dumas [Inria, Senior Researcher, until August 2017, now External Collaborator]
Georges Gonthier [Inria, Senior Researcher]
Pierre Lairez [Inria, Researcher]
Assia Mahboubi [Team co-leader, Inria, Researcher, until September 2017]

External Collaborator

Marc Mezzarobba [CNRS]

Technical Staff

Maxence Guesdon [Inria, Engineer, 40%, until May 2017]

PhD Student

Thomas Sibut-Pinote [École Polytechnique]

Administrative Assistant

Christine Biard [Inria, until Aug 2017]

2. Overall Objectives

2.1. Scientific challenges, expected impact

The general orientation of our team is described by the short name given to it: *Special Functions*, that is, particular mathematical functions that have established names due to their importance in mathematical analysis, physics, and other application domains. Indeed, we ambition to study special functions with the computer, by combined means of computer algebra and formal methods.

Computer-algebra systems have been advertised for decades as software for "doing mathematics by computer" [65]. For instance, computer-algebra libraries can uniformly generate a corpus of mathematical properties about special functions, so as to display them on an interactive website. This possibility was recently shown by the computer-algebra component of the team [19]. Such an automated generation significantly increases the reliability of the mathematical corpus, in comparison to the content of existing static authoritative handbooks. The importance of the validity of these contents can be measured by the very wide audience that such handbooks have had, to the point that a book like [14] remains one of the most cited mathematical publications ever and has motivated the 10-year-long project of writing its successor [16]. However, can the mathematics produced "by computer" be considered as *true* mathematics? More specifically, whereas it is nowadays well established that the computer helps in discovering and observing new mathematical phenomenons, can the mathematical statements produced with the aid of the computer and the mathematical results computed by it be accepted as valid mathematics, that is, as having the status of mathematical *proofs*? Beyond the reported weaknesses or controversial design choices of mainstream computer-algebra systems, the issue is more of an epistemological nature. It will not find its solution even in the advent of the ultimate computer-algebra system: the social process of peer-reviewing just falls short of evaluating the results produced by computers, as reported by Th. Hales [44] after the publication of his proof of the Kepler Conjecture about sphere packing.

A natural answer to this deadlock is to move to an alternative kind of mathematical software and to use a proof assistant to check the correctness of the desired properties or formulas. The success of large-scale formalization projects, like the Four-Color Theorem of graph theory [39], the above-mentioned Kepler Conjecture [44], and the Odd Order Theorem of group theory [0], have increased the understanding of the appropriate software-engineering methods for this peculiar kind of programming. For computer algebra, this legitimates a move to proof assistants now.

The Dynamic Dictionary of Mathematical Functions [0] (DDMF) [19] is an online computer-generated handbook of mathematical functions that ambitions to serve as a reference for a broad range of applications. This software was developed by the computer-algebra component of the team as a project [0] of the MSR–INRIA Joint Centre. It bases on a library for the computer-algebra system Maple, Algolib [0], whose development started 20 years ago in ÉPI Algorithms [0]. As suggested by the constant questioning of certainty by new potential users, DDMF deserves a formal guarantee of correctness of its content, on a level that proof assistants can provide. Fortunately, the maturity of special-functions algorithms in Algolib makes DDMF a stepping stone for such a formalization: it provides a well-understood and unified algorithmic treatment, without which a formal certification would simply be unreachable.

The formal-proofs component of the team emanates from another project of the MSR–INRIA Joint Centre, namely the Mathematical Components project (MathComp) [0]. Since 2006, the MathComp group has endeavoured to develop computer-checked libraries of formalized mathematics, using the Coq proof assistant [61]. The methodological aim of the project was to understand the design methods leading to successful large-scale formalizations. The work culminated in 2012 with the completion of a formal proof of the Odd Order Theorem, resulting in the largest corpus of algebraic theories ever machine-checked with a proof assistant and a whole methodology to effectively combine these components in order to tackle complex formalizations. In particular, these libraries provide a good number of the many algebraic objects needed to reason about special functions and their properties, like rational numbers, iterated sums, polynomials, and a rich hierarchy of algebraic structures.

The present team takes benefit from these recent advances to explore the formal certification of the results collected in DDMF. The aim of this project is to concentrate the formalization effort on this delimited area, building on DDMF and the Algolib library, as well as on the Coq system [61] and on the libraries developed by the MathComp project.

[0] https://www.msr-inria.fr/news/the-formalization-of-the-odd-order-theorem-has-been-completed-the-20-septembre-2012/
[0] http://ddmf.msr-inria.inria.fr/1.9.1/ddmf
[0] https://www.msr-inria.fr/projects/dynamic-dictionary-of-mathematical-functions/
[0] http://algo.inria.fr/libraries/
[0] http://algo.inria.fr/
[0] http://www.msr-inria.fr/projects/mathematical-components/

2.1.1. Use computer algebra but convince users beyond reasonable doubt

The following few opinions on computer algebra are, we believe, typical of computer-algebra users' doubts and difficulties when using computer-algebra systems:

- Fredrik Johansson, expert in the multi-precision numerical evaluation of special functions and in fast computer-algebra algorithms, writes on his blog [50]: "Mathematica is great for cross-checking numerical values, but it's not unusual to run into bugs, so *triple checking is a good habit*." One answer in the discussion is: "We can claim that Mathematica has [...] *an impossible to understand semantics*: If Mathematica's output is wrong then change the input. If you don't like the answer, change the question. That seems to be the philosophy behind."
- A professor's advice to students [57] on using Maple: "You may wish to use Maple to check your homework answers. If you do then keep in mind that Maple sometimes gives the *wrong answer, usually because you asked incorrectly, or because of niceties of analytic continuation*. You may even be bitten by an occasional Maple bug, though that has become fairly unlikely. Even with as powerful a tool as Maple you will still *have to devise your own checks* and you will still have to think."
- Jacques Carette, former head of the maths group at Maplesoft, about a bug [15] when asking Maple to take the limit `limit(f(n) * exp(-n), n = infinity)` for an undetermined function `f`: "The problem is that there is an *implicit assumption in the implementation* that unknown functions do not 'grow too fast'."

As explained by the expert views above, complaints by computer-algebra users are often due to their misunderstanding of what a computer-algebra systems is, namely a purely syntactic tool for calculations, that the user must complement with a semantics. Still, robustness and consistency of computer-algebra systems are not ensured as of today, and, whatever Zeilberger may provocatively say in his Opinion 94 [66], a firmer logical foundation is necessary. Indeed, the fact is that many "bugs" in a computer-algebra system cannot be fixed by just the usual debugging method of tracking down the faulty lines in the code. It is sort of "by design": assumptions that too often remain implicit are really needed by the design of symbolic algorithms and cannot easily be expressed in the programming languages used in computer algebra. A similar certification initiative has already been undertaken in the domain of numerical computing, in a successful manner [48], [22]. It is natural to undertake a similar approach for computer algebra.

2.1.2. Make computer algebra and formal proofs help one another

Some of the mathematical objects that interest our team are still totally untouched by formalization. When implementing them and their theory inside a proof assistant, we have to deal with the pervasive discrepancy between the published literature and the actual implementation of computer-algebra algorithms. Interestingly, this forces us to clarify our computer-algebraic view on them, and possibly make us discover holes lurking in published (human) proofs. We are therefore convinced that the close interaction of researchers from both fields, which is what we strive to maintain in this team, is a strong asset.

For a concrete example, the core of Zeilberger's creative telescoping manipulates rational functions up to simplifications. In summation applications, checking that these simplifications do not hide problematic divisions by 0 is most often left to the reader. In the same vein, in the case of integrals, the published algorithms do not check the convergence of all integrals, especially in intermediate calculations. Such checks are again left to the readers. In general, we expect to revisit the existing algorithms to ensure that they are meaningful for genuine mathematical sequences or functions, and not only for algebraic idealizations.

Another big challenge in this project originates in the scientific difference between computer algebra and formal proofs. Computer algebra seeks speed of calculation on *concrete instances* of algebraic data structures (polynomials, matrices, etc). For their part, formal proofs manipulate symbolic expressions in terms of *abstract variables* understood to represent generic elements of algebraic data structures. In view of this, a continuous challenge is to develop the right, hybrid thinking attitude that is able to effectively manage concrete and abstract values simultaneously, alternatively computing and proving with them.

2.1.3. Experimental mathematics with special functions

Applications in combinatorics and mathematical physics frequently involve equations of so high orders and so large sizes, that computing or even storing all their coefficients is impossible on existing computers. Making this tractable is an extraordinary challenge. The approach we believe in is to design algorithms of good—ideally quasi-optimal—complexity in order to extract precisely the required data from the equations, while avoiding the computationally intractable task of completely expanding them into an explicit representation.

Typical applications with expected high impact are the automatic discovery and algorithmic proof of results in combinatorics and mathematical physics for which human proofs are currently unattainable.

2.2. Research axes

The implementation of certified symbolic computations on special functions in the Coq proof assistant requires both investigating new formalization techniques and renewing the traditional computer-algebra viewpoint on these standard objects. Large mathematical objects typical of computer algebra occur during formalization, which also requires us to improve the efficiency and ergonomics of Coq. In order to feed this interdisciplinary activity with new motivating problems, we additionally pursue a research activity oriented towards experimental mathematics in application domains that involve special functions. We expect these applications to pose new algorithmic challenges to computer algebra, which in turn will deserve a formal-certification effort. Finally, DDMF is the motivation and the showcase of our progress on the certification of these computations. While striving to provide a formal guarantee of the correctness of the information it displays, we remain keen on enriching its mathematical content by developing new computer-algebra algorithms.

2.2.1. Computer algebra certified by the Coq system

Our formalization effort consists in organizing a cooperation between a computer-algebra system and a proof assistant. The computer-algebra system is used to produce efficiently algebraic data, which are later processed by the proof assistant. The success of this cooperation relies on the design of appropriate libraries of formalized mathematics, including certified implementations of certain computer-algebra algorithms. On the other side, we expect that scrutinizing the implementation and the output of computer-algebra algorithms will shed a new light on their semantics and on their correctness proofs, and help clarifying their documentation.

2.2.1.1. Libraries of formalized mathematics

The appropriate framework for the study of efficient algorithms for special functions is *algebraic*. Representing algebraic theories as Coq formal libraries takes benefit from the methodology emerging from the success of ambitious projects like the formal proof of a major classification result in finite-group theory (the Odd Order Theorem) [37].

Yet, a number of the objects we need to formalize in the present context has never been investigated using any interactive proof assistant, despite being considered as commonplaces in computer algebra. For instance there is up to our knowledge no available formalization of the theory of non-commutative rings, of the algorithmic theory of special-functions closures, or of the asymptotic study of special functions. We expect our future formal libraries to prove broadly reusable in later formalizations of seemingly unrelated theories.

2.2.1.2. Manipulation of large algebraic data in a proof assistant

Another peculiarity of the mathematical objects we are going to manipulate with the Coq system is their size. In order to provide a formal guarantee on the data displayed by DDMF, two related axes of research have to be pursued. First, efficient algorithms dealing with these large objects have to be programmed and run in Coq. Recent evolutions of the Coq system to improve the efficiency of its internal computations [17], [20] make this objective reachable. Still, how to combine the aforementioned formalization methodology with these cutting-edge evolutions of Coq remains one of the prospective aspects of our project. A second need is to help users *interactively* manipulate large expressions occurring in their conjectures, an objective for which little has been done so far. To address this need, we work on improving the ergonomics of the system in two ways:

first, ameliorating the reactivity of Coq in its interaction with the user; second, designing and implementing extensions of its interface to ease our formalization activity. We expect the outcome of these lines of research to be useful to a wider audience, interested in manipulating large formulas on topics possibly unrelated to special functions.

2.2.1.3. *Formal-proof-producing normalization algorithms*

Our algorithm certifications inside Coq intend to simulate well-identified components of our Maple packages, possibly by reproducing them in Coq. It would however not have been judicious to re-implement them inside Coq in a systematic way. Indeed for a number of its components, the output of the algorithm is more easily checked than found, like for instance the solving of a linear system. Rather, we delegate the discovery of the solutions to an external, untrusted oracle like Maple. Trusted computations inside Coq then formally validate the correctness of the a priori untrusted output. More often than not, this validation consists in implementing and executing normalization procedures *inside* Coq. A challenge of this automation is to make sure they go to scale while remaining efficient, which requires a Coq version of non-trivial computer-algebra algorithms. A first, archetypal example we expect to work on is a non-commutative generalization of the normalization procedure for elements of rings [43].

2.2.2. Better symbolic computations with special functions

Generally speaking, we design algorithms for manipulating special functions symbolically, whether univariate or with parameters, and for extracting algorithmically any kind of algebraic and analytic information from them, notably asymptotic properties. Beyond this, the heart of our research is concerned with parametrised definite summations and integrations. These very expressive operations have far-ranging applications, for instance, to the computation of integral transforms (Laplace, Fourier) or to the solution of combinatorial problems expressed via integrals (coefficient extractions, diagonals). The algorithms that we design for them need to really operate on the level of linear functional systems, differential and of recurrence. In all cases, we strive to design our algorithms with the constant goal of good theoretical complexity, and we observe that our algorithms are also fast in practice.

2.2.2.1. *Special-function integration and summation*

Our long-term goal is to design fast algorithms for a general method for special-function integration (*creative telescoping*), and make them applicable to general special-function inputs. Still, our strategy is to proceed with simpler, more specific classes first (rational functions, then algebraic functions, hyperexponential functions, D-finite functions, non-D-finite functions; two variables, then many variables); as well, we isolate analytic questions by first considering types of integration with a more purely algebraic flavor (constant terms, algebraic residues, diagonals of combinatorics). In particular, we expect to extend our recent approach [25] to more general classes (algebraic with nested radicals, for example): the idea is to speed up calculations by making use of an analogue of Hermite reduction that avoids considering certificates. Homologous problems for summation will be addressed as well.

2.2.2.2. *Applications to experimental mathematics*

As a consequence of our complexity-driven approach to algorithms design, the algorithms mentioned in the previous paragraph are of good complexity. Therefore, they naturally help us deal with applications that involve equations of high orders and large sizes.

With regard to combinatorics, we expect to advance the algorithmic classification of combinatorial classes like walks and urns. Here, the goal is to determine if enumerative generating functions are rational, algebraic, or D-finite, for example. Physical problems whose modelling involves special-function integrals comprise the study of models of statistical mechanics, like the Ising model for ferro-magnetism, or questions related to Hamiltonian systems.

Number theory is another promising domain of applications. Here, we attempt an experimental approach to the automated certification of integrality of the coefficients of mirror maps for Calabi–Yau manifolds. This could also involve the discovery of new Calabi–Yau operators and the certification of the existing ones. We also plan to algorithmically discover and certify new recurrences yielding good approximants needed in irrationality proofs.

It is to be noted that in all of these application domains, we would so far use general algorithms, as was done in earlier works of ours [24], [29], [27]. To push the scale of applications further, we plan to consider in each case the specifics of the application domain to tailor our algorithms.

2.2.3. Interactive and certified mathematical web sites

In continuation of our past project of an encyclopedia at http://ddmf.msr-inria.inria.fr/1.9.1/ddmf, we ambition to both enrich and certify the formulas about the special functions that we provide online. For each function, our website shows its essential properties and the mathematical objects attached to it, which are often infinite in nature (numerical evaluations, asymptotic expansions). An interactive presentation has the advantage of allowing for adaption to the user's needs. More advanced content will broaden the encyclopedia:

- the algorithmic discussion of equations with parameters, leading to certified automatic case analysis based on arithmetic properties of the parameters;
- lists of summation and integral formulas involving special functions, including validity conditions on the parameters;
- guaranteed large-precision numerical evaluations.

3. Research Program

3.1. Studying special functions by computer algebra

Computer algebra manipulates symbolic representations of exact mathematical objects in a computer, in order to perform computations and operations like simplifying expressions and solving equations for "closed-form expressions". The manipulations are often fundamentally of algebraic nature, even when the ultimate goal is analytic. The issue of efficiency is a particular one in computer algebra, owing to the extreme swell of the intermediate values during calculations.

Our view on the domain is that research on the algorithmic manipulation of special functions is anchored between two paradigms:

- adopting linear differential equations as the right data structure for special functions,
- designing efficient algorithms in a complexity-driven way.

It aims at four kinds of algorithmic goals:

- algorithms combining functions,
- functional equations solving,
- multi-precision numerical evaluations,
- guessing heuristics.

This interacts with three domains of research:

- computer algebra, meant as the search for quasi-optimal algorithms for exact algebraic objects,
- symbolic analysis/algebraic analysis;
- experimental mathematics (combinatorics, mathematical physics, ...).

This view is made explicit in the present section.

3.1.1. Equations as a data structure

Numerous special functions satisfy linear differential and/or recurrence equations. Under a mild technical condition, the existence of such equations induces a finiteness property that makes the main properties of the functions decidable. We thus speak of *D-finite functions*. For example, 60 % of the chapters in the handbook [14] describe D-finite functions. In addition, the class is closed under a rich set of algebraic operations. This makes linear functional equations just the right data structure to encode and manipulate special functions. The power of this representation was observed in the early 1990s [67], leading to the design of many algorithms in computer algebra. Both on the theoretical and algorithmic sides, the study of D-finite functions shares much with neighbouring mathematical domains: differential algebra, D-module theory, differential Galois theory, as well as their counterparts for recurrence equations.

3.1.2. Algorithms combining functions

Differential/recurrence equations that define special functions can be recombined [67] to define: additions and products of special functions; compositions of special functions; integrals and sums involving special functions. Zeilberger's fast algorithm for obtaining recurrences satisfied by parametrised binomial sums was developed in the early 1990s already [68]. It is the basis of all modern definite summation and integration algorithms. The theory was made fully rigorous and algorithmic in later works, mostly by a group in RISC (Linz, Austria) and by members of the team [56], [64], [33], [30], [31], [51]. The past ÉPI Algorithms contributed several implementations (*gfun* [59], *Mgfun* [33]).

3.1.3. Solving functional equations

Encoding special functions as defining linear functional equations postpones some of the difficulty of the problems to a delayed solving of equations. But at the same time, solving (for special classes of functions) is a sub-task of many algorithms on special functions, especially so when solving in terms of polynomial or rational functions. A lot of work has been done in this direction in the 1990s; more intensively since the 2000s, solving differential and recurrence equations in terms of special functions has also been investigated.

3.1.4. Multi-precision numerical evaluation

A major conceptual and algorithmic difference exists for numerical calculations between data structures that fit on a machine word and data structures of arbitrary length, that is, *multi-precision* arithmetic. When multi-precision floating-point numbers became available, early works on the evaluation of special functions were just promising that "most" digits in the output were correct, and performed by heuristically increasing precision during intermediate calculations, without intended rigour. The original theory has evolved in a twofold way since the 1990s: by making computable all constants hidden in asymptotic approximations, it became possible to guarantee a *prescribed* absolute precision; by employing state-of-the-art algorithms on polynomials, matrices, etc, it became possible to have evaluation algorithms in a time complexity that is linear in the output size, with a constant that is not more than a few units. On the implementation side, several original works exist, one of which (*NumGfun* [55]) is used in our DDMF.

3.1.5. Guessing heuristics

"Differential approximation", or "Guessing", is an operation to get an ODE likely to be satisfied by a given approximate series expansion of an unknown function. This has been used at least since the 1970s and is a key stone in spectacular applications in experimental mathematics [29]. All this is based on subtle algorithms for Hermite–Padé approximants [18]. Moreover, guessing can at times be complemented by proven quantitative results that turn the heuristics into an algorithm [26]. This is a promising algorithmic approach that deserves more attention than it has received so far.

3.1.6. Complexity-driven design of algorithms

The main concern of computer algebra has long been to prove the feasibility of a given problem, that is, to show the existence of an algorithmic solution for it. However, with the advent of faster and faster computers, complexity results have ceased to be of theoretical interest only. Nowadays, a large track of works in computer algebra is interested in developing fast algorithms, with time complexity as close as possible to linear in their output size. After most of the more pervasive objects like integers, polynomials, and matrices have been endowed with fast algorithms for the main operations on them [38], the community, including ourselves, started to turn its attention to differential and recurrence objects in the 2000s. The subject is still not as developed as in the commutative case, and a major challenge remains to understand the combinatorics behind summation and integration. On the methodological side, several paradigms occur repeatedly in fast algorithms: "divide and conquer" to balance calculations, "evaluation and interpolation" to avoid intermediate swell of data, etc. [23].

3.2. Trusted computer-algebra calculations

3.2.1. Encyclopedias

Handbooks collecting mathematical properties aim at serving as reference, therefore trusted, documents. The decision of several authors or maintainers of such knowledge bases to move from paper books [14], [16], [60] to websites and wikis [0] allows for a more collaborative effort in proof reading. Another step toward further confidence is to manage to generate the content of an encyclopedia by computer-algebra programs, as is the case with the Wolfram Functions Site [0] or DDMF [0]. Yet, due to the lingering doubts about computer-algebra systems, some encyclopedias propose both cross-checking by different systems and handwritten companion paper proofs of their content [0]. As of today, there is no encyclopedia certified with formal proofs.

3.2.2. Computer algebra and symbolic logic

Several attempts have been made in order to extend existing computer-algebra systems with symbolic manipulations of logical formulas. Yet, these works are more about extending the expressivity of computer-algebra systems than about improving the standards of correctness and semantics of the systems. Conversely, several projects have addressed the communication of a proof system with a computer-algebra system, resulting in an increased automation available in the proof system, to the price of the uncertainty of the computations performed by this oracle.

3.2.3. Certifying systems for computer algebra

More ambitious projects have tried to design a new computer-algebra system providing an environment where the user could both program efficiently and elaborate formal and machine-checked proofs of correctness, by calling a general-purpose proof assistant like the Coq system. This approach requires a huge manpower and a daunting effort in order to re-implement a complete computer-algebra system, as well as the libraries of formal mathematics required by such formal proofs.

3.2.4. Semantics for computer algebra

The move to machine-checked proofs of the mathematical correctness of the output of computer-algebra implementations demands a prior clarification about the often implicit assumptions on which the presumably correctly implemented algorithms rely. Interestingly, this preliminary work, which could be considered as independent from a formal certification project, is seldom precise or even available in the literature.

3.2.5. Formal proofs for symbolic components of computer-algebra systems

A number of authors have investigated ways to organize the communication of a chosen computer-algebra system with a chosen proof assistant in order to certify specific components of the computer-algebra systems, experimenting various combinations of systems and various formats for mathematical exchanges. Another line of research consists in the implementation and certification of computer-algebra algorithms inside the logic [63], [43], [52] or as a proof-automation strategy. Normalization algorithms are of special interest when they allow to check results possibly obtained by an external computer-algebra oracle [36]. A discussion about the systematic separation of the search for a solution and the checking of the solution is already clearly outlined in [49].

3.2.6. Formal proofs for numerical components of computer-algebra systems

Significant progress has been made in the certification of numerical applications by formal proofs. Libraries formalizing and implementing floating-point arithmetic as well as large numbers and arbitrary-precision arithmetic are available. These libraries are used to certify floating-point programs, implementations of mathematical functions and for applications like hybrid systems.

[0] for instance http://dlmf.nist.gov/ for special functions or http://oeis.org/ for integer sequences
[0] http://functions.wolfram.com/
[0] http://ddmf.msr-inria.inria.fr/1.9.1/ddmf
[0] http://129.81.170.14/~vhm/Table.html

3.3. Machine-checked proofs of formalized mathematics

To be checked by a machine, a proof needs to be expressed in a constrained, relatively simple formal language. Proof assistants provide facilities to write proofs in such languages. But, as merely writing, even in a formal language, does not constitute a formal proof just per se, proof assistants also provide a proof checker: a small and well-understood piece of software in charge of verifying the correctness of arbitrarily large proofs. The gap between the low-level formal language a machine can check and the sophistication of an average page of mathematics is conspicuous and unavoidable. Proof assistants try to bridge this gap by offering facilities, like notations or automation, to support convenient formalization methodologies. Indeed, many aspects, from the logical foundation to the user interface, play an important role in the feasibility of formalized mathematics inside a proof assistant.

3.3.1. Logical foundations and proof assistants

While many logical foundations for mathematics have been proposed, studied, and implemented, type theory is the one that has been more successfully employed to formalize mathematics, to the notable exception of the Mizar system [53], which is based on set theory. In particular, the calculus of construction (CoC) [34] and its extension with inductive types (CIC) [35], have been studied for more than 20 years and been implemented by several independent tools (like Lego, Matita, and Agda). Its reference implementation, Coq [61], has been used for several large-scale formalizations projects (formal certification of a compiler back-end; four-color theorem). Improving the type theory underlying the Coq system remains an active area of research. Other systems based on different type theories do exist and, whilst being more oriented toward software verification, have been also used to verify results of mainstream mathematics (prime-number theorem; Kepler conjecture).

3.3.2. Computations in formal proofs

The most distinguishing feature of CoC is that computation is promoted to the status of rigorous logical argument. Moreover, in its extension CIC, we can recognize the key ingredients of a functional programming language like inductive types, pattern matching, and recursive functions. Indeed, one can program effectively inside tools based on CIC like Coq. This possibility has paved the way to many effective formalization techniques that were essential to the most impressive formalizations made in CIC.

Another milestone in the promotion of the computations-as-proofs feature of Coq has been the integration of compilation techniques in the system to speed up evaluation. Coq can now run realistic programs in the logic, and hence easily incorporates calculations into proofs that demand heavy computational steps.

Because of their different choice for the underlying logic, other proof assistants have to simulate computations outside the formal system, and indeed fewer attempts to formalize mathematical proofs involving heavy calculations have been made in these tools. The only notable exception, which was finished in 2014, the Kepler conjecture, required a significant work to optimize the rewriting engine that simulates evaluation in Isabelle/HOL.

3.3.3. Large-scale computations for proofs inside the Coq system

Programs run and proved correct inside the logic are especially useful for the conception of automated decision procedures. To this end, inductive types are used as an internal language for the description of mathematical objects by their syntax, thus enabling programs to reason and compute by case analysis and recursion on symbolic expressions.

The output of complex and optimized programs external to the proof assistant can also be stamped with a formal proof of correctness when their result is easier to *check* than to *find*. In that case one can benefit from their efficiency without compromising the level of confidence on their output at the price of writing and certify a checker inside the logic. This approach, which has been successfully used in various contexts, is very relevant to the present research project.

3.3.4. Relevant contributions from the Mathematical Component libraries

Representing abstract algebra in a proof assistant has been studied for long. The libraries developed by the MathComp project for the proof of the Odd Order Theorem provide a rather comprehensive hierarchy of structures; however, they originally feature a large number of instances of structures that they need to organize. On the methodological side, this hierarchy is an incarnation of an original work [37] based on various mechanisms, primarily type inference, typically employed in the area of programming languages. A large amount of information that is implicit in handwritten proofs, and that must become explicit at formalization time, can be systematically recovered following this methodology.

Small-scale reflection [40] is another methodology promoted by the MathComp project. Its ultimate goal is to ease formal proofs by systematically dealing with as many bureaucratic steps as possible, by automated computation. For instance, as opposed to the style advocated by Coq's standard library, decidable predicates are systematically represented using computable boolean functions: comparison on integers is expressed as program, and to state that $a \leq b$ one compares the output of this program run on a and b with $true$. In many cases, for example when a and b are values, one can prove or disprove the inequality by pure computation.

The MathComp library was consistently designed after uniform principles of software engineering. These principles range from simple ones, like naming conventions, to more advanced ones, like generic programming, resulting in a robust and reusable collection of formal mathematical components. This large body of formalized mathematics covers a broad panel of algebraic theories, including of course advanced topics of finite group theory, but also linear algebra, commutative algebra, Galois theory, and representation theory. We refer the interested reader to the online documentation of these libraries [62], which represent about 150,000 lines of code and include roughly 4,000 definitions and 13,000 theorems.

Topics not addressed by these libraries and that might be relevant to the present project include real analysis and differential equations. The most advanced work of formalization on these domains is available in the HOL-Light system [45], [46], [47], although some existing developments of interest [21], [54] are also available for Coq. Another aspect of the MathComp libraries that needs improvement, owing to the size of the data we manipulate, is the connection with efficient data structures and implementations, which only starts to be explored.

3.3.5. User interaction with the proof assistant

The user of a proof assistant describes the proof he wants to formalize in the system using a textual language. Depending on the peculiarities of the formal system and the applicative domain, different proof languages have been developed. Some proof assistants promote the use of a declarative language, when the Coq and Matita systems are more oriented toward a procedural style.

The development of the large, consistent body of MathComp libraries has prompted the need to design an alternative and coherent language extension for the Coq proof assistant [42], [41], enforcing the robustness of proof scripts to the numerous changes induced by code refactoring and enhancing the support for the methodology of small-scale reflection.

The development of large libraries is quite a novelty for the Coq system. In particular any long-term development process requires the iteration of many refactoring steps and very little support is provided by most proof assistants, with the notable exception of Mizar [58]. For the Coq system, this is an active area of research.

4. Highlights of the Year

4.1. Highlights of the Year

4.1.1. Awards

Pierre Lairez was awarded the SIAM/AAG (SIAM Activity Group on Algebraic Geometry) Early Career Prize.

5. New Software and Platforms

5.1. DynaMoW

Dynamic Mathematics on the Web

FUNCTIONAL DESCRIPTION: Programming tool for controlling the generation of mathematical websites that embed dynamical mathematical contents generated by computer-algebra calculations. Implemented in OCaml.

- Participants: Alexis Darrasse, Frédéric Chyzak and Maxence Guesdon
- Contact: Frédéric Chyzak
- URL: http://ddmf.msr-inria.inria.fr/DynaMoW/

5.2. ECS

Encyclopedia of Combinatorial Structures

FUNCTIONAL DESCRIPTION: On-line mathematical encyclopedia with an emphasis on sequences that arise in the context of decomposable combinatorial structures, with the possibility to search by the first terms in the sequence, keyword, generating function, or closed form.

- Participants: Alexis Darrasse, Frédéric Chyzak, Maxence Guesdon and Stéphanie Petit
- Contact: Frédéric Chyzak
- URL: http://ecs.inria.fr/

5.3. DDMF

Dynamic Dictionary of Mathematical Functions

FUNCTIONAL DESCRIPTION: Web site consisting of interactive tables of mathematical formulas on elementary and special functions. The formulas are automatically generated by OCaml and computer-algebra routines. Users can ask for more terms of the expansions, more digits of the numerical values, proofs of some of the formulas, etc.

- Participants: Alexandre Benoit, Alexis Darrasse, Bruno Salvy, Christoph Koutschan, Frédéric Chyzak, Marc Mezzarobba, Maxence Guesdon, Stefan Gerhold and Thomas Gregoire
- Contact: Frédéric Chyzak
- URL: http://ddmf.msr-inria.inria.fr/1.9.1/ddmf

5.4. Mgfun

multivariate generating functions package

FUNCTIONAL DESCRIPTION: The Mgfun Project is a collection of packages for the computer algebra system Maple, and is intended for the symbolic manipulation of a large class of special functions and combinatorial sequences (in one or several variables and indices) that appear in many branches of mathematics, mathematical physics, and engineering sciences. Members of the class satisfy a crucial finiteness property which makes the class amenable to computer algebra methods and enjoy numerous algorithmic closure properties, including algorithmic closures under integration and summation.

- Contact: Frédéric Chyzak
- URL: http://specfun.inria.fr/chyzak/mgfun.html

5.5. Ssreflect

FUNCTIONAL DESCRIPTION: Ssreflect is a tactic language extension to the Coq system, developed by the Mathematical Components team.

- Participants: Assia Mahboubi, Cyril Cohen, Enrico Tassi, Georges Gonthier, Laurence Rideau, Laurent Théry and Yves Bertot
- Contact: Yves Bertot
- URL: http://math-comp.github.io/math-comp/

5.6. Math-Components

Mathematical Components library
FUNCTIONAL DESCRIPTION: The Mathematical Components library is a set of Coq libraries that cover the mechanization of the proof of the Odd Order Theorem.

RELEASE FUNCTIONAL DESCRIPTION: The library includes 16 more theory files, covering in particular field and Galois theory, advanced character theory, and a construction of algebraic numbers.

- Participants: Alexey Solovyev, Andrea Asperti, Assia Mahboubi, Cyril Cohen, Enrico Tassi, François Garillot, Georges Gonthier, Ioana Pasca, Jeremy Avigad, Laurence Rideau, Laurent Théry, Russell O'Connor, Sidi Ould Biha, Stéphane Le Roux and Yves Bertot
- Contact: Assia Mahboubi
- URL: http://math-comp.github.io/math-comp/

5.7. CoqInterval

Interval package for Coq
KEYWORDS: Interval arithmetic - Coq
FUNCTIONAL DESCRIPTION: CoqInterval is a library for the proof assistant Coq.

It provides several tactics for proving theorems on enclosures of real-valued expressions. The proofs are performed by an interval kernel which relies on a computable formalization of floating-point arithmetic in Coq.

The Marelle team developed a formalization of rigorous polynomial approximation using Taylor models in Coq. In 2014, this library has been included in CoqInterval.

- Participants: Assia Mahboubi, Érik Martin-Dorel, Guillaume Melquiond, Jean-Michel Muller, Laurence Rideau, Laurent Théry, Micaela Mayero, Mioara Joldes, Nicolas Brisebarre and Thomas Sibut-Pinote
- Contact: Guillaume Melquiond
- Publications: Proving bounds on real-valued functions with computations - Floating-point arithmetic in the Coq system - Proving Tight Bounds on Univariate Expressions with Elementary Functions in Coq - Formally Verified Approximations of Definite Integrals - Formally Verified Approximations of Definite Integrals
- URL: http://coq-interval.gforge.inria.fr/

6. New Results

6.1. Efficient Algorithms in Computer Algebra

This year has seen the end of the writing and the publication of a book on computer-algebra algorithms [8]. The course at Master 2 level *Algorithmes efficaces en calcul formel* is a course that Alin Bostan and Frédéric Chyzak have set up progressively since 2005 together with Marc Giusti (LIX), Bruno Salvy (today AriC), as well as, initially, Éric Schost (LIX at the time) and François Ollivier (LIX), and, more recently, Grégoire Lecerf (LIX). The course is very strongly focused to presenting the design of algorithms guided by complexity analysis, with the goal to lead the students to the understanding of all algorithmic aspects that are necessary to the "creative telescoping" used for symbolic computations of sums and integrals. Their lecture notes had been circulating in and used by the (French) computer-algebra community, while they long had the goal of turning them into a book. They could publish it in 2017 (686 pages), after a big finalization effort in 2016 and 2017. The first parts of the book present fast algorithms for basic objects (integers, polynomials, series, matrices, linear recurrences), insisting on general principles to design efficient algorithms. The next parts of the work build on them to address topics that have made recent progress: factorization of polynomials, algorithms for polynomial systems, definite summation and integration. The work [8] is online as a HAL collection[0]. It is available for free in pdf format and is otherwise sold at a very low price (via print-on-demand). Over the first three months after publication, the book has sold roughly 60 printed copies and the pdf has been downloaded 265 times.

6.2. Hypergeometric Expressions for Generating Functions of Walks with Small Steps in the Quarter Plane

In [2], Alin Bostan and Frédéric Chyzak, together with Mark van Hoeij (Florida State University), Manuel Kauers (Johannes Kepler University), and Lucien Pech, have studied nearest-neighbors walks on the two-dimensional square lattice, that is, models of walks on $\mathbb{Z}^2$ defined by a fixed step set that consists of non-zero vectors with coordinates 0, 1 or -1. They concerned themselves with the enumeration of such walks starting at the origin and constrained to remain in the quarter plane $\mathbb{N}^2$, counted by their length and by the position of their ending point. In earlier works, Bousquet-Mélou and Mishna had identified 19 models of walks that possess a D-finite generating function, and linear differential equations had then been guessed in these cases by Bostan and Kauers. Here, we have given the first proof that these equations are indeed satisfied by the corresponding generating functions. As a first corollary, we have proved that all these 19 generating functions can be expressed in terms of Gauss' hypergeometric functions, with specific parameters that relate them intimately to elliptic integrals. As a second corollary, we have shown that all the 19 generating functions are transcendental, and that among their 19×4 combinatorially meaningful specializations only four are algebraic functions.

6.3. Multiple Binomial Sums

Multiple binomial sums form a large class of multi-indexed sequences, closed under partial summation, which contains most of the sequences obtained by multiple summation of products of binomial coefficients, as well as all the sequences with algebraic generating function. Alin Bostan and Pierre Lairez, together with Bruno Salvy (AriC), have studied in [7] the representation of the generating functions of binomial sums by integrals of rational functions. The outcome is twofold. Firstly, we have shown that a univariate sequence is a multiple binomial sum if and only if its generating function is the diagonal of a rational function. Secondly, we have proposed algorithms that decide the equality of multiple binomial sums and that compute recurrence relations for them. In conjunction with geometric simplifications of the integral representations, this approach behaves well in practice. The process avoids the computation of certificates and the problem of the appearance of spurious singularities that afflicts discrete creative telescoping, both in theory and in practice.

[0] https://hal.archives-ouvertes.fr/AECF/

6.4. Algebraic Diagonals and Walks

The diagonal of a multivariate power series F is the univariate power series $\mathsf{Diag}F$ generated by the diagonal terms of F. Diagonals form an important class of power series; they occur frequently in number theory, theoretical physics and enumerative combinatorics. In [28], Alin Bostan and Louis Dumont, together with Bruno Salvy (AriC), have studied algorithmic questions related to diagonals in the case where F is the Taylor expansion of a bivariate rational function. It is classical that in this case $\mathsf{Diag}F$ is an algebraic function. They have proposed an algorithm for computing an annihilating polynomial of $\mathsf{Diag}F$. They have given a precise bound on the size of this polynomial and show that generically, this polynomial is the minimal polynomial of $\mathsf{Diag}F$ and that its size reaches the bound. Their algorithm runs in time quasi-linear in this bound, which grows exponentially with the degree of the input rational function. They have also addressed the related problem of enumerating directed lattice walks. The insight given by their study has led to a new method for expanding the generating power series of bridges, excursions and meanders. They have shown that their first N terms can be computed in quasi-linear complexity in N, without first computing a very large polynomial equation. An extended version of this work has been presented in [4].

6.5. A Human Proof of the Gessel Conjecture

Counting lattice paths obeying various geometric constraints is a classical topic in combinatorics and probability theory. Many recent works deal with the enumeration of 2-dimensional walks with prescribed steps confined to the positive quadrant. A notoriously difficult case concerns the so-called *Gessel walks*: they are planar walks confined to the positive quarter plane, which move by unit steps in any of the West, North-East, East, and South-West directions. In 2001, Ira Gessel conjectured a closed-form expression for the number of such walks of a given length starting and ending at the origin. In 2008, Kauers, Koutschan and Zeilberger gave a computer-aided proof of this conjecture. The same year, Bostan and Kauers showed, using again computer algebra tools, that the trivariate generating function of Gessel walks is algebraic. This year, Alin Bostan, together with Irina Kurkova (Univ. Paris 6) and Kilian Raschel (CNRS and Univ. Tours), proposed in [6] the first “human proofs” of these results. They are derived from a new expression for the generating function of Gessel walks in terms of special functions.

6.6. Subresultants in Multiple Roots

In [3], we have provided explicit formulae for the coefficients of the order-d polynomial subresultant of $(x-\alpha)^m$ and $(x-\beta)^n$ with respect to the set of Bernstein polynomials $\{(x-\alpha)^j(x-\beta)^{d-j}, 0 \le j \le d\}$. They are given by hypergeometric expressions arising from determinants of binomial Hankel matrices.

6.7. On Matrices with Displacement Structure: Generalized Operators and Faster Algorithms

For matrices with displacement structure, basic operations like multiplication, inversion, and linear-system solving can all be expressed in terms of a single task: evaluating the product AB, where A is a structured $n \times n$ matrix of displacement rank α, and B is an arbitrary $n \times \alpha$ matrix. Given B and a so-called *generator* of A, this product is classically computed with a cost ranging from $O(\alpha^2 \mathsf{M}(n))$ to $O(\alpha^2 \mathsf{M}(n) \log(n))$ arithmetic operations, depending on the specific structure of A. (Here, M is a cost function for polynomial multiplication.) In [5], Alin Bostan, jointly with Claude-Pierre Jeannerod (AriC), Christophe Mouilleron (ENSIIE), and Éric Schost (University of Waterloo), has generalized classical displacement operators, based on block diagonal matrices with companion diagonal blocks, and has also designed fast algorithms to perform the task above for this extended class of structured matrices. The cost of these algorithms ranges from $O(\alpha^{\omega-1}\mathsf{M}(n))$ to $O(\alpha^{\omega-1}\mathsf{M}(n)\log(n))$, with ω such that two $n \times n$ matrices over a field can be multiplied using $O(n^\omega)$ field operations. By combining this result with classical randomized regularization techniques, he has obtained faster Las Vegas algorithms for structured inversion and linear system solving.

6.8. Quasilinear Average Complexity for Solving Polynomial Systems

How many operations do we need on the average to compute an approximate root of a random Gaussian polynomial system? Beyond Smale's 17th problem that asked whether a polynomial bound is possible, Pierre Lairez has proved in [10] a quasi-optimal bound (input size)$^{1+o(1)}$, which improves upon the previously known (input size)$^{3/2+o(1)}$ bound. His new algorithm relies on numerical continuation along *rigid continuation paths*. The central idea is to consider rigid motions of the equations rather than line segments in the linear space of all polynomial systems. This leads to a better average condition number and allows for bigger steps. He showed that on the average, one approximate root of a random Gaussian polynomial system of n equations of degree at most D in $n+1$ homogeneous variables can be computed with $O(n^5D^2)$ continuation steps. This is a decisive improvement over previous bounds, which prove no better than $\sqrt{2}^{\min(n,D)}$ continuation steps on the average.

6.9. Computing the Homology of Basic Semialgebraic Sets in Weak Exponential Time

In [9], Pierre Lairez, jointly with Peter Bürgisser (TU Berlin) and Felipe Cucker (City University of Hong Kong), has described and analyzed an algorithm for computing the homology (Betti numbers and torsion coefficients) of basic semialgebraic sets. The algorithm works in weak exponential time, that is, out of a set of exponentially small measure in the space of data, the cost of the algorithm is exponential in the size of the data. All algorithms previously proposed for this problem have a complexity that is doubly exponential (and this is so for almost all data).

6.10. Formally Certified Computation of Improper Definite Integrals

Assia Mahboubi and Thomas Sibut-Pinote, in collaboration with Guillaume Melquiond (Toccata), have pursued their work on the certified computation of intervals approximating the values of definite integrals involving elementary mathematical functions. This library provides an automated tool that builds a formal proof of the correctness of the output, that is, a formal proof that the interval contains the mathematical values and a formal proof of the integrability of the input function on the input interval. This tool has been extended this year, and it can now deal with improper integrals, that is, integrals whose bounds are infinite or singularities of the integrand. The methodology, the implementation and benchmarks have been described in [13].

6.11. A Complete Formal Proof of the Irrationality of $\zeta(3)$

Assia Mahboubi and Thomas Sibut-Pinote have completed a formal proof of the irrationality of the constant $\zeta(3)$. The missing step in a previous work [32] with Frédéric Chyzak and Enrico Tassi was to obtain a formal proof of the asymptotic behaviour of the least common multiple of the first n integers. They have written a report on this work, which is included as a chapter in Thomas Sibut-Pinote's PhD manuscript.

7. Partnerships and Cooperations

7.1. National Initiatives

7.1.1. ANR

FastRelax (ANR-14-CE25-0018). Goal: Develop computer-aided proofs of numerical values, with certified and reasonably tight error bounds, without sacrificing efficiency. Leader: B. Salvy (Inria, ENS Lyon). Participants: Assia Mahboubi, Th. Sibut-Pinote. Website: http://fastrelax.gforge.inria.fr/.

7.2. International Research Visitors

7.2.1. Visits of International Scientists

- Marni Mishna (Simon Fraser University) visited the team for one week in January.
- Emre Sertöz (Max Planck Institute Leipzig) visited the team for one week in November. He worked with Pierre Lairez on applications to algebraic geometry of two tools developped at Specfun: the computations of periods (Lairez's PhD) and numerical analytic continuation (Mezzarobba's PhD, 2011).
- Karen Yeats (Simon Fraser University) visited the team for a few days in June. She continued a work on bijective combinatorics of words with Frédéric Chyzak. A text is now under writing.

7.2.1.1. Internships

- Pascal Fong did a Master internship from March to August. Under the supervision of Pierre Lairez and Mohab Safey El Din (UPMC), he studied the numerical computation of the length of plane algebraic curves.
- Rémy Garnier did a Master internship from March to July. Under the supervision of Alin Bostan and Frédéric Chyzak, he studied existing algorithms to solve linear differential systems for their rational-function solutions.
- Meissa M'baye did a Master internship from February to June. Under the remote supervision of Assia Mahboubi, he studied the principles of proof assistants and surveyed formalization methodologies for elementary number theory.

7.2.2. Visits to International Teams

- Frédéric Chyzak and Alin Bostan have been invited by the Erwin Schrödinger Institute (Vienna, Austria) for two weeks, to participate to the thematic program "Algorithmic and Enumerative Combinatorics" http://www.mat.univie.ac.at/~kratt/esi4/.
- Pierre Lairez visited Felipe Cucker (City University of Hong Kong) for two weeks. The outcome is a strengthened collaboration on the study of the complexity of numerical algorithms. A publication is in preparation: the second part of [10].
- Georges Gonthier was invited at the Newton Institue, for six weeks, as co-organiser and participant to the Big Proof thematic program.
- Assia Mahboubi visited Sander Dahmen (VU Amsterdam, The Netherlands) for three days. She has started a collaboration with his team, to obtain formal guarantees of computations for number theory.
- Assia Mahboubi has been invited by the Newton Institute (Cambridge, UK) for one month. She participated to the Big Proof thematic program.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events Organisation

8.1.1.1. General Chair, Scientific Chair

- Alin Bostan is part of the Scientific advisory board of the conference series *Effective Methods in Algebraic Geometry* (MEGA).
- Frédéric Chyzak is member of the steering committee of the *Journées Nationales de Calcul Formel* (JNCF), the annual meeting of the French computer algebra community.

- Frédéric Chyzak is elected member of the steering committee of the *International Symposium on Symbolic and Algebraic Computation* (ISSAC, 3-year term, 2016–2018).
- Assia Mahboubi has served on the scientific committee of the Journées Scientifiques Inria and of the EUTypes summer school.

8.1.1.2. Member of the Organizing Committees

- Georges Gonthier was co-chair or the organising committee of the Big Proof thematic program held at the Newton Institute (Cambridge, UK) in June-August 2017.
- Assia Mahboubi has organized the TTT workshop, satellite of the POPL'17 conference.

8.1.1.3. Other

The team organizes a regular seminar[0], with roughly 15–20 talks a year. The topics reflect the team's interests: computer algebra, combinatorics, number theory, formal proofs, and related domains.

8.1.2. Scientific Events Selection

8.1.2.1. Member of the Conference Program Committees

- Assia Mahboubi has served as member of the conference program committees of the international conferences CPP 17, CADE 17, ITP 17, and on the program committee of the international workshops TYPES 17, TyDe 17, HoTT/UF 17.

8.1.2.2. Reviewer

- Alin Bostan has served as reviewer for the selection of the international conferences ISSAC 2017 and MEGA 2017.
- Frédéric Chyzak has served as reviewer for the selection of the international conference ISSAC 2017.
- Assia Mahboubi has served as external reviewer for the international conference LICS 17.

8.1.3. Journal

8.1.3.1. Member of the Editorial Boards

- Georges Gonthier is on the editorial board of the Journal of Formalized Reasoning.

8.1.3.2. Reviewer - Reviewing Activities

- Alin Bostan has served as a reviewer for the journals: *Journal of Symbolic Computation*; *Linear Algebra and its Applications*; *Journal of Algebra and its Applications*; *Journal of Complexity*; *Advances in Applied Mathematics*; *Journal of Combinatorial Theory, Series A*.
- Assia Mahboubi has served as reviewer for the journals: *Annals of Mathematics and Artificial Intelligence*; *Journal of Automated Reasoning*.

8.1.4. Invited Talks

- Alin Bostan has been invited to give a series of three lectures at the *JNCF – Journées Nationales de Calcul Formel* (CIRM, Luminy, France), January 16–18, 2017, http://jncf2017.lip6.fr.
- Alin Bostan has been invited to give a talk at the workshop *EDATE – Equations différentielles : aspects théoriques et effectifs*, Grenoble, March 13–15, 2017, http://edate2017.sciencesconf.org/.
- Alin Bostan has been invited to give a talk at the workshop *ANT – Automata in Number Theory*, île de Porquerolles, May 30—June 2, 2017, http://indico.math.cnrs.fr/event/2347/.
- Alin Bostan has been invited to give a talk at the workshop *Lattice walks at the Interface of Algebra, Analysis and Combinatorics*, BIRS, Banff, Canada, September 17—22, 2017, http://www.birs.ca/events/2017/5-day-workshops/17w5090.

[0] https://specfun.inria.fr/seminar/

- Alin Bostan has been invited to give a series of introductory lectures at the *Workshop on Computer Algebra in Combinatorics*, Erwin Schrödinger Institut (ESI), Vienna, Austria, November 13–17, 2017, http://www.mat.univie.ac.at/~kratt/esi4/workshop2.html.
- Frédéric Chyzak has been invited to give a talk at the *Second International Conference "Computer Algebra in Moscow"*, Plekhanov Russian University of Economics, Moscow, Russia, October 30 to November 30, 2017, http://www.ccas.ru/ca/conference.
- Frédéric Chyzak has been invited to give a talk at the *Workshop on Computer Algebra in Combinatorics*, Erwin Schrödinger Institut (ESI), Vienna, Austria, November 13–17, 2017, http://www.mat.univie.ac.at/~kratt/esi4/workshop2.html.
- Georges Gonthier gave an invited talk at the Special Session on Computer-Aided Proofs of the *Association for Symbolic Logic 2017 North American Meeting*, Boise, Idaho, USA, March 20–23, 2017.
- Georges Gonthier gave an invited talk at the *Second Conference on Artificial Intelligence and Theorem Proving (AITP'17)*, Obergürgl, Austria, March 26–30, 2017.
- Georges Gonthier gave a talk at the *ERCIM Workshop on Blockchains*, Paris, May 23, 2017.
- Georges Gonthier gave an invited talk at the Workshop *Computer-aided Mathematical Proof*, part of the Big Proof Program, Isaac Newton Institute, Cambridge, U.K.
- Pierre Lairez gave an invited talk at *Effective Methods in Algebraic Geometry (MEGGA 2017)*, Nice.
- Pierre Lairez gave an invited talk at the *Conference on Foundations of Computational Mathematics (FoCM 2017)*, Barcelona, Spain.
- Assia Mahboubi has been invited to give a talk at the General Mathematics Colloquium of the VU Amsterdam, The Netherlands.
- Assia Mahboubi has been invited to give a talk at the Workshop *Computer-aided Mathematical Proof*, part of the Big Proof Program, Isaac Newton Institute, Cambridge, U.K.

8.1.5. Leadership within the Scientific Community

- Assia Mahboubi leads the working group *Type theory based tools* inside the EUTYPES COST project. She is also a member of the management committee for France for this project and a member of its core management group.

8.1.6. Research Administration

- Georges Gonthier serves on the Conseil de l'École Doctorale de Mathématiques Hadamard.
- Assia Mahboubi has been a member of the *Commission Scientifique* of Inria Saclay — Île-de-France, until September 2017.

8.2. Teaching - Supervision - Juries

- Alin Bostan has served as a jury member of the French *Agrégation de Mathématiques – épreuve de modélisation, option C*.

8.2.1. Teaching

Licence:

Thomas Sibut-Pinote, *Les bases de la programmation et de l'algorithmique*, 32h, L3, École polytechnique, France.

Thomas Sibut-Pinote, *Les principes des langages de programmation*, 32h, L3, École polytechnique, France.

Master:

Frédéric Chyzak, *Algorithmes efficaces en calcul formel*, 18h, M2, MPRI, France.

Alin Bostan, *Algorithmes efficaces en calcul formel*, 40.5h, M2, MPRI, France.

Pierre Lairez, *Algorithmique avancée*, 18h, M1, École polytechnique, France.

Assia Mahboubi, *Algorithmes d'élimination des quantificateurs*, 3h, M2, Université Rennes 1, France.

8.2.2. Supervision

HdR : Alin Bostan, *Computer algebra for lattice path combinatorics* [1], Université Paris 13, December 15, 2017.

PhD : Thomas Sibut-Pinote, *Investigations en Mathématiques Assistées par Ordinateur: Expérimentation, Calcul et Certification*, Université Paris-Saclay, December 4, 2017.

8.2.3. Juries

- Frédéric Chyzak has served as an examiner in the PhD jury of Cyril Hugounenq *Volcans et calcul d'isogénies*, Université de Versailles – Saint-Quentin-en-Yvelines, September 25, 2017.
- Frédéric Chyzak has been a member of the hiring jury at Inria (Concours CR 2017).
- Georges Gonthier served on the Habilitation à diriger des Recherches of Paul-André Melliès *Une étude micrologique de la négation*, Université Paris Diderot, November 20, 2017.
- Assia Mahboubi has been a member of the hiring jury at Inria (Concours CR 2017).
- Assia Mahboubi has served as an examiner in the PhD jury of Evmorfia-Iro Bartzia *Une formalisation des courbes elliptiques pour la cryptographie*, Université Paris-Saclay, February 15, 2017.
- Assia Mahboubi has served as an examiner in the PhD jury of Étienne Miquey *Réalisabilité classique et effets de bord*, Université Paris Diderot, November 17, 2017.
- Assia Mahboubi has been a member of the hiring jury of a Maître de conférence position at Université Paris Diderot.

8.3. Popularization

- Assia Mahboubi has written an article for the MathExpress journal, at the occasion of the *salon Culture & Jeux Mathématiques*. See the Maths Language express volume at http://www.cijm.org/accueil/productions-cijm/90-maths-express.

9. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] A. Bostan.*Computer Algebra for Lattice path Combinatorics*, Université Paris 13, December 2017, Habilitation à diriger des recherches, https://hal.archives-ouvertes.fr/tel-01660300.

Articles in International Peer-Reviewed Journal

[2] A. Bostan, F. Chyzak, M. van Hoeij, M. Kauers, L. Pech.*Hypergeometric Expressions for Generating Functions of Walks with Small Steps in the Quarter Plane*, in "European Journal of Combinatorics", 2017, vol. 61, p. 242-275, https://arxiv.org/abs/1606.02982 [*DOI :* 10.1016/j.ejc.2016.10.010], https://hal.inria.fr/hal-01332175.

[3] A. Bostan, C. D'Andrea, T. Krick, A. Szanto, M. Valdettaro.*Subresultants in multiple roots: an extremal case*, in "Linear Algebra and its Applications", 2017, vol. 529, p. 185-198, https://arxiv.org/abs/1608.03740 [*DOI :* 10.1016/j.laa.2017.04.019], https://hal.archives-ouvertes.fr/hal-01588546.

[4] A. BOSTAN, L. DUMONT, B. SALVY.*Algebraic Diagonals and Walks: Algorithms, Bounds, Complexity*, in "Journal of Symbolic Computation", 2017, vol. 83, p. 68–92, https://arxiv.org/abs/1510.04526 [*DOI :* 10.1016/J.JSC.2016.11.006], https://hal.archives-ouvertes.fr/hal-01244914.

[5] A. BOSTAN, C.-P. JEANNEROD, C. MOUILLERON, E. SCHOST.*On Matrices With Displacement Structure: Generalized Operators and Faster Algorithms*, in "SIAM Journal on Matrix Analysis and Applications", 2017, vol. 38, n^{o} 3, p. 733-775, https://arxiv.org/abs/1703.03734 [*DOI :* 10.1137/16M1062855], https://hal.archives-ouvertes.fr/hal-01588552.

[6] A. BOSTAN, I. KURKOVA, K. RASCHEL.*A human proof of Gessel's lattice path conjecture*, in "Transactions of the American Mathematical Society", 2017, vol. 369, n^{o} 2, February 2017, p. 1365-1393, https://arxiv.org/abs/1309.1023 - Published electronically: April 14, 2016, https://hal.archives-ouvertes.fr/hal-00858083.

[7] A. BOSTAN, P. LAIREZ, B. SALVY.*Multiple binomial sums*, in "Journal of Symbolic Computation", 2017, vol. 80, n^{o} 2, p. 351–386 [*DOI :* 10.1016/J.JSC.2016.04.002], https://hal.archives-ouvertes.fr/hal-01220573.

Scientific Books (or Scientific Book chapters)

[8] A. BOSTAN, F. CHYZAK, M. GIUSTI, R. LEBRETON, G. LECERF, B. SALVY, E. SCHOST.*Algorithmes Efficaces en Calcul Formel*, published by the Authors, 2017, Voir la page du livre à l'adresse https://hal.archives-ouvertes.fr/AECF/, https://hal.inria.fr/hal-01431717.

Other Publications

[9] P. BÜRGISSER, F. CUCKER, P. LAIREZ.*Computing the Homology of Basic Semialgebraic Sets in Weak Exponential Time*, June 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01545657.

[10] P. LAIREZ.*Rigid continuation paths I. Quasilinear average complexity for solving polynomial systems*, November 2017, working paper or preprint, https://hal.inria.fr/hal-01631778.

[11] H. LOMBARDI, A. MAHBOUBI.*Théories géométriques pour l'algèbre des nombres réels*, April 2017, working paper or preprint, https://hal.inria.fr/hal-01426164.

[12] S. MADDAH, M. A. BARKATOU.*Formal Solutions of Singularly Perturbed Linear Differential Systems*, January 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01417265.

[13] A. MAHBOUBI, G. MELQUIOND, T. SIBUT-PINOTE.*Formally Verified Approximations of Definite Integrals*, February 2017, working paper or preprint, https://hal.inria.fr/hal-01630143.

References in notes

[14] M. ABRAMOWITZ, I. A. STEGUN (editors). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, Dover, New York, 1992, xiv+1046, Reprint of the 1972 edition.

[15] *Computer Algebra Errors*, Article in mathematics blog MathOverflow, http://mathoverflow.net/questions/11517/computer-algebra-errors.

[16] F. W. J. OLVER, D. W. LOZIER, R. F. BOISVERT, C. W. CLARK (editors). *NIST Handbook of mathematical functions*, Cambridge University Press, 2010.

[17] M. ARMAND, B. GRÉGOIRE, A. SPIWACK, L. THÉRY.*Extending Coq with Imperative Features and its Application to SAT Verication*, in "Interactive Theorem Proving, international Conference, ITP 2010, Edinburgh, Scotland, July 11–14, 2010, Proceedings", Lecture Notes in Computer Science, Springer, 2010.

[18] B. BECKERMANN, G. LABAHN.*A uniform approach for the fast computation of matrix-type Padé approximants*, in "SIAM J. Matrix Anal. Appl.", 1994, vol. 15, n^{o} 3, p. 804–823.

[19] A. BENOIT, F. CHYZAK, A. DARRASSE, S. GERHOLD, M. MEZZAROBBA, B. SALVY.*The Dynamic Dictionary of Mathematical Functions (DDMF)*, in "The Third International Congress on Mathematical Software (ICMS 2010)", K. FUKUDA, J. VAN DER HOEVEN, M. JOSWIG, N. TAKAYAMA (editors), Lecture Notes in Computer Science, 2010, vol. 6327, p. 35–41, http://dx.doi.org/10.1007/978-3-642-15582-6_7.

[20] M. BOESPFLUG, M. DÉNÈS, B. GRÉGOIRE.*Full reduction at full throttle*, in "First International Conference on Certified Programs and Proofs, Taiwan, December 7–9", Lecture Notes in Computer Science, Springer, 2011.

[21] S. BOLDO, C. LELAY, G. MELQUIOND.*Improving Real Analysis in Coq: A User-Friendly Approach to Integrals and Derivatives*, in "Certified Programs and Proofs", C. HAWBLITZEL, D. MILLER (editors), Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2012, vol. 7679, p. 289-304, http://dx.doi.org/10.1007/978-3-642-35308-6_22.

[22] S. BOLDO, G. MELQUIOND.*Flocq: A Unified Library for Proving Floating-point Algorithms in Coq*, in "Proceedings of the 20th IEEE Symposium on Computer Arithmetic", Tübingen, Germany, July 2011, p. 243–252.

[23] A. BOSTAN.*Algorithmes rapides pour les polynômes, séries formelles et matrices*, in "Actes des Journées Nationales de Calcul Formel", Luminy, France, 2010, p. 75–262, Les cours du CIRM, tome 1, numéro 2, http://ccirm.cedram.org:80/ccirm-bin/fitem?id=CCIRM_2010__1_2_75_0.

[24] A. BOSTAN, S. BOUKRAA, S. HASSANI, J.-M. MAILLARD, J.-A. WEIL, N. ZENINE.*Globally nilpotent differential operators and the square Ising model*, in "J. Phys. A: Math. Theor.", 2009, vol. 42, n^{o} 12, 50, http://dx.doi.org/10.1088/1751-8113/42/12/125206.

[25] A. BOSTAN, S. CHEN, F. CHYZAK, Z. LI.*Complexity of creative telescoping for bivariate rational functions*, in "ISSAC'10: Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation", New York, NY, USA, ACM, 2010, p. 203–210, http://doi.acm.org/10.1145/1837934.1837975.

[26] A. BOSTAN, F. CHYZAK, G. LECERF, B. SALVY, É. SCHOST.*Differential equations for algebraic functions*, in "ISSAC'07: Proceedings of the 2007 international symposium on Symbolic and algebraic computation", C. W. BROWN (editor), ACM Press, 2007, p. 25–32, http://dx.doi.org/10.1145/1277548.1277553.

[27] A. BOSTAN, F. CHYZAK, M. VAN HOEIJ, L. PECH.*Explicit formula for the generating series of diagonal 3D rook paths*, in "Sém. Loth. Comb.", 2011, vol. B66a, 27, http://www.emis.de/journals/SLC/wpapers/s66bochhope.html.

[28] A. BOSTAN, L. DUMONT, B. SALVY.*Algebraic Diagonals and Walks*, in "ISSAC'15 International Symposium on Symbolic and Algebraic Computation", Bath, United Kingdom, ACM Press, July 2015, p. 77–84 [*DOI :* 10.1145/2755996.2756663], https://hal.archives-ouvertes.fr/hal-01240729.

[29] A. BOSTAN, M. KAUERS.*The complete generating function for Gessel walks is algebraic*, in "Proceedings of the American Mathematical Society", September 2010, vol. 138, n° 9, p. 3063–3078, With an appendix by Mark van Hoeij.

[30] F. CHYZAK.*An extension of Zeilberger's fast algorithm to general holonomic functions*, in "Discrete Math.", 2000, vol. 217, n° 1-3, p. 115–134, Formal power series and algebraic combinatorics (Vienna, 1997).

[31] F. CHYZAK, M. KAUERS, B. SALVY.*A Non-Holonomic Systems Approach to Special Function Identities*, in "ISSAC'09: Proceedings of the Twenty-Second International Symposium on Symbolic and Algebraic Computation", J. MAY (editor), 2009, p. 111–118, http://dx.doi.org/10.1145/1576702.1576720.

[32] F. CHYZAK, A. MAHBOUBI, T. SIBUT-PINOTE, E. TASSI.*A Computer-Algebra-Based Formal Proof of the Irrationality of $\zeta(3)$*, in "ITP - 5th International Conference on Interactive Theorem Proving", Vienna, Austria, 2014, https://hal.inria.fr/hal-00984057.

[33] F. CHYZAK, B. SALVY.*Non-commutative elimination in Ore algebras proves multivariate identities*, in "J. Symbolic Comput.", 1998, vol. 26, n° 2, p. 187–227.

[34] T. COQUAND, G. P. HUET.*The Calculus of Constructions*, in "Inf. Comput.", 1988, vol. 76, n° 2/3, p. 95-120, http://dx.doi.org/10.1016/0890-5401(88)90005-3.

[35] T. COQUAND, C. PAULIN-MOHRING.*Inductively defined types*, in "Proceedings of Colog'88", P. MARTIN-LÖF, G. MINTS (editors), Lecture Notes in Computer Science, Springer-Verlag, 1990, vol. 417.

[36] D. DELAHAYE, M. MAYERO.*Dealing with algebraic expressions over a field in Coq using Maple*, in "J. Symbolic Comput.", 2005, vol. 39, n° 5, p. 569–592, Special issue on the integration of automated reasoning and computer algebra systems, http://dx.doi.org/10.1016/j.jsc.2004.12.004.

[37] F. GARILLOT, G. GONTHIER, A. MAHBOUBI, L. RIDEAU.*Packaging Mathematical Structures*, in "Theorem Proving in Higher-Order Logics", S. BERGHOFER, T. NIPKOW, C. URBAN, M. WENZEL (editors), Lecture Notes in Computer Science, Springer, 2009, vol. 5674, p. 327–342.

[38] J. VON ZUR. GATHEN, J. GERHARD.*Modern computer algebra*, 2nd, Cambridge University Press, New York, 2003, xiv+785.

[39] G. GONTHIER.*Formal proofs—the four-colour theorem*, in "Notices of the AMS", 2008, vol. 55, n° 11, p. 1382-1393.

[40] G. GONTHIER, A. MAHBOUBI.*An introduction to small scale reflection in Coq*, in "Journal of Formalized Reasoning", 2010, vol. 3, n° 2, p. 95–152.

[41] G. GONTHIER, A. MAHBOUBI, E. TASSI.*A Small Scale Reflection Extension for the Coq system*, Inria, 2008, n° RR-6455, http://hal.inria.fr/inria-00258384.

[42] G. GONTHIER, E. TASSI.*A language of patterns for subterm selection*, in "ITP", LNCS, 2012, vol. 7406, p. 361–376.

[43] B. GRÉGOIRE, A. MAHBOUBI.*Proving Equalities in a Commutative Ring Done Right in Coq*, in "Theorem Proving in Higher Order Logics, 18th International Conference, TPHOLs 2005, Oxford, UK, August 22-25, 2005, Proceedings", Lecture Notes in Computer Science, Springer, 2005, vol. 3603, p. 98–113.

[44] T. HALES.*Formal proof*, in "Notices of the AMS", 2008, vol. 55, n° 11, p. 1370-1380.

[45] J. HARRISON.*A HOL Theory of Euclidean space*, in "Theorem Proving in Higher Order Logics, 18th International Conference, TPHOLs 2005", Oxford, UK, J. HURD, T. MELHAM (editors), Lecture Notes in Computer Science, Springer-Verlag, 2005, vol. 3603.

[46] J. HARRISON.*Formalizing an analytic proof of the prime number theorem*, in "Journal of Automated Reasoning", 2009, vol. 43, p. 243–261, Dedicated to Mike Gordon on the occasion of his 60th birthday.

[47] J. HARRISON.*Theorem proving with the real numbers*, CPHC/BCS distinguished dissertations, Springer, 1998.

[48] J. HARRISON.*A Machine-Checked Theory of Floating Point Arithmetic*, in "Theorem Proving in Higher Order Logics: 12th International Conference, TPHOLs'99", Nice, France, Y. BERTOT, G. DOWEK, A. HIRSCHOWITZ, C. PAULIN, L. THÉRY (editors), Lecture Notes in Computer Science, Springer-Verlag, 1999, vol. 1690, p. 113–130.

[49] J. HARRISON, L. THÉRY.*A Skeptic's Approach to Combining HOL and Maple*, in "J. Autom. Reason.", December 1998, vol. 21, n° 3, p. 279–294, http://dx.doi.org/10.1023/A:1006023127567.

[50] F. JOHANSSON.*Another Mathematica bug*, Article on personal blog, http://fredrik-j.blogspot.fr/2009/07/another-mathematica-bug.html.

[51] C. KOUTSCHAN.*A fast approach to creative telescoping*, in "Math. Comput. Sci.", 2010, vol. 4, n° 2-3, p. 259–266, http://dx.doi.org/10.1007/s11786-010-0055-0.

[52] A. MAHBOUBI.*Implementing the cylindrical algebraic decomposition within the Coq system*, in "Mathematical Structures in Computer Science", 2007, vol. 17, n° 1, p. 99–127.

[53] R. MATUSZEWSKI, P. RUDNICKI.*Mizar: the first 30 years*, in "Mechanized Mathematics and Its Applications", 2005, vol. 4.

[54] M. MAYERO.*Problèmes critiques et preuves formelles*, Université Paris 13, novembre 2012, Habilitation à Diriger des Recherches.

[55] M. MEZZAROBBA.*NumGfun: a package for numerical and analytic computation and D-finite functions*, in "ISSAC 2010—Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation", New York, ACM, 2010, p. 139–146, http://dx.doi.org/10.1145/1837934.1837965.

[56] P. PAULE, M. SCHORN.*A Mathematica version of Zeilberger's algorithm for proving binomial coefficient identities*, in "J. Symbolic Comput.", 1995, vol. 20, n° 5-6, p. 673–698, Symbolic computation in combinatorics Δ_1 (Ithaca, NY, 1993), http://dx.doi.org/10.1006/jsco.1995.1071.

[57] B. PETERSEN.*Maple*, Personal web site.

[58] P. RUDNICKI, A. TRYBULEC. *On the Integrity of a Repository of Formalized Mathematics*, in "Proceedings of the Second International Conference on Mathematical Knowledge Management", London, UK, MKM '03, Springer-Verlag, 2003, p. 162–174, http://dl.acm.org/citation.cfm?id=648071.748518.

[59] B. SALVY, P. ZIMMERMANN. *Gfun: a Maple package for the manipulation of generating and holonomic functions in one variable*, in "ACM Trans. Math. Software", 1994, vol. 20, n^{o} 2, p. 163–177.

[60] N. J. A. SLOANE, S. PLOUFFE. *The Encyclopedia of Integer Sequences*, Academic Press, San Diego, 1995.

[61] THE COQ DEVELOPMENT TEAM. *The Coq Proof Assistant: Reference Manual*, http://coq.inria.fr/doc/.

[62] THE MATHEMATICAL COMPONENT TEAM. *A Formalization of the Odd Order Theorem using the Coq proof assistant*, September 2012, http://www.msr-inria.fr/projects/mathematical-components/.

[63] L. THÉRY. *A Machine-Checked Implementation of Buchberger's Algorithm*, in "J. Autom. Reasoning", 2001, vol. 26, n^{o} 2, p. 107-137, http://dx.doi.org/10.1023/A:1026518331905.

[64] K. WEGSCHAIDER. *Computer generated proofs of binomial multi-sum identities*, RISC, J. Kepler University, May 1997, 99.

[65] S. WOLFRAM. *Mathematica: A system for doing mathematics by computer (2nd ed.)*, Addison-Wesley, 1992, I.

[66] D. ZEILBERGER. *Opinion 94: The Human Obsession With "Formal Proofs" is a Waste of the Computer's Time, and, Even More Regretfully, of Humans' Time*, 2009, http://www.math.rutgers.edu/~zeilberg/Opinion94.html.

[67] D. ZEILBERGER. *A holonomic systems approach to special functions identities*, in "J. Comput. Appl. Math.", 1990, vol. 32, n^{o} 3, p. 321–368.

[68] D. ZEILBERGER. *The method of creative telescoping*, in "J. Symbolic Comput.", 1991, vol. 11, n^{o} 3, p. 195–204.

Team TAU

TAckling the Underspecified

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions).

RESEARCH CENTER
Saclay - Île-de-France

THEME
Optimization, machine learning and statistical methods

Table of contents

Team TAU

Creation of the Team: 2016 December 01

Keywords:

Computer Science and Digital Science:

A3.3.3. - Big data analysis
A3.4. - Machine learning and statistics
A3.5.2. - Recommendation systems
A8.2. - Optimization
A8.6. - Information theory
A9.2. - Machine learning
A9.3. - Signal analysis

Other Research Topics and Application Domains:

B4. - Energy
B7.2.1. - Smart vehicles
B9.1.2. - Serious games
B9.4.3. - Physics
B9.4.5. - Data science
B9.5.10. - Digital humanities

1. Personnel

Research Scientists

Marc Schoenauer [Team leader, Inria, Senior Researcher, HDR]
Guillaume Charpiat [Inria, Researcher]
Cyril Furtlehner [Inria, Researcher]
Yann Ollivier [CNRS, Researcher, until May 2017, HDR]
Martine Sebag [CNRS, Senior Researcher, HDR]
Paola Tubaro [CNRS, Researcher]

Faculty Members

Cécile Germain [Univ. Paris-Sud, Professor, HDR]
Isabelle Guyon [Univ. Paris-Sud + Inria chair, Professor + Advanced Research Position]
Philippe Caillou [Univ Paris-Sud, Associate Professor]
Aurélien Decelle [Univ. Paris-Sud, Associate Professor]

External Collaborators

Bernat Gacias [Other Univ., until Mar 2017]
Flora Jay [CNRS]
Balasz Kegl [CNRS]
Nilo Schwenke [Ecole Polytechnique]

Technical Staff

Yasmina Bouzbiba [Inria, from Apr 2017]
Etienne Brame [Inria]
Raphaël Jaiswal [Inria, from Sep 2017]
Felix Louistisserand [Inria]
Luis Marti Orosa [Inria]

Dejan Tusar [Inria, until Sep 2017]

PhD Students

Nacim Belkhir [Thalès, until May 2017]
Victor Berger [Inria, from Oct 2017]
Vincent Berthier [Inria, until Nov 2017]
Benjamin Donnot [RTE]
Guillaume Doquet [Univ Paris-Sud]
Victor Estrade [Univ Paris-Sud]
Giancarlo Fissore [Inria, from Oct 2017]
François Gonard [Institut de recherche technologique System X]
Diviyan Kalainathan [Univ Paris-Sud]
Zhengying Liu [Univ Paris-Sud (AMX), from Oct 2017]
Pierre-Yves Masse [Inria, until Dec. 2017]
Marc Nabhan [Renault, from Jun 2017]
Adrian Pol [Organisation européenne pour la recherche nucléaire]
Herilalaina Rakotoarison [Inria, from Nov 2017]
Théophile Sanchez [Univ Paris-Sud, from Oct 2017]
Thomas Schmitt [Inria]
Lisheng Sun [Univ Paris-Sud]
Corentin Tallec [Univ. Paris-Sud (AMX)]
Aris Tritas [Inria, Intern from Apr-Sep 2017, PhD since Oct. 2017]
Pierre Wolinski [Univ. Paris-Sud]

Post-Doctoral Fellows

Berna Bakir Batu [Inria]
Edgar Galvan Lopez [Programme Marie-Curie, until Mar 2017]
Olivier Goudet [Inria]

Visiting Scientist

Mandar Chandorkar [CWI, from Sep 2017]

Administrative Assistant

Olga Mwana Mobulakani [Inria, until May 2017 – no assistant Jun-Dec 2017]

2. Overall Objectives

2.1. Presentation

The last two years have been a turning point for the team. Of course, since its creation in 2003, TAO activities had constantly but slowly evolved, as old problems were being solved, and new applications arose. But recent abrupt progresses in Machine Learning (and in particular in Deep Learning) have greatly accelerated these changes also within the team. It so happened that this change of slope also coincided with some more practical changes in TAO ecosystem: following Inria 12-years rule, the team definitely ended in December 2016. The new team TAU (for **TA**ckling the **U**nderspecified) has been proposed, and the creation process is on-going. At the same time important staff changes took place, that also justify even sharper changes in the team focus. The year 2017 is hence the first year of a new era for the (remaining) members of the team, and some research topics might still continue to change in the next months.

As said above, several permanent members left the team. The good news is that Anne Auger and Nikolaus Hansen left TAO only to create, together with Dimo Brockoff (from the Dolphin team in Lille) their own offspring team, RANDOPT, also in the Saclay Inria Center.

The departure of both Olivier Teytaud and Yann Ollivier is a not-so-good news, in the sense that this participates to the general brain drain that is observed in all French academia toward the large private AI research centers: Olivier Teytaud (since June 2016) and Yann Ollivier (since May 2017) are now with respectively Google Zurich and Facebook AI Research in Paris. Whereas Yann continues his collaboration with us, and in particular still the supervision of his current PhD students (with more to come), Olivier's activities in TAO had to be continued by other permament researchers, namely Isabelle Guyon (see below) and Marc Schoenauer (who took over the last months of the PhD students supervisions - with Olivier in the background). We nevertheless have started a new collaboration with Olivier.

Finally, to end this presentation of the evolution of the team personal on a positive note, we were lucky enough that Isabelle Guyon joined the team (as of January 2016), as full Professor at Université Paris-Sud, and on the Inria/Université Paris-Sud Chair on Data Science for E-Science. She contributed not only by bringing her own research themes to the team, but also by enthusiastically taking over most activities of the team in the Energy application domain (see Section 4.1), the main context of Olivier's work in the recent years.

Next Sections will describe in more detail the current positionning of the new TAU team.

2.2. Context and overall goal of the project

Building upon the expertise in machine learning (ML) and stochastic optimization of the late TAO project-team, the TAU team aims to tackle **the vagueness of the Big Data purposes**. Based on the claim that (sufficiently) big data can to some extent compensate for the lack of knowledge, Big Data is hoped to fulfill all Artificial Intelligence commitments. This makes Big Data under-specified in three respects:

- A first source of under-specification is related to **common sense**, and the gap between observation and interpretation. The acquired data do not report on "obvious" issues; still, obvious issues are not necessarily so for the computer. Providing the machine with common sense is a many-faceted, AI long, challenge. A current challenge is to **interpret the data** and cope with its blind zones.
- A second source of under-specification regards the **steering of a Big Data system**. Such systems commonly require constant learning in order to deal with open environments and users with diverse profiles, expertise and expectations. A Big Data system thus is a dynamic process, whose behavior will depend in a cumulative way upon its future environment. The question regards the **control of a lifelong learning system**.
- A third source of under-specification regards its social acceptability. There is little doubt that Big Data can pave the way for Big Brother, and ruin the social contract through modeling benefits and costs at the individual level. What are the **fair trade-offs between safety, freedom and efficiency** ? We do not know the answers. A first practical and scientific challenge is to assess the fairness of a solution.

The tackling of the under-specified issues in Big Data in TAU currently relies on four core research dimensions, taking inspiration and validation in four main application domains. These research dimensions involve Causal Modelling (required to support prescriptive Big Data), Deep Learning (related to constructive representations, and their compositionality), Optimization and Meta-Optimization (including sequential decision making and categorization of problems), and Big-Data Driven Design. The application domains include the long-lasting domains of Energy Management and High Energy Physics, the more recent focus of TAO/TAU in Computational Social and Economic Sciences, and, new this year, the Autonomous Vehicle, and Population Genetics.

3. Research Program

3.1. Causal modelling

Data science is viewed as an information processing cycle: i) exploiting data and prior knowledge to build models; ii) using models to support optimal decisions in view of desirable ends; and iii) acquiring more data in order to refine the models and/or the desirable ends. Inasmuch data science supports prescriptive recommendations, it requires building causal models: these hold in front of interventions on the application domain – as opposed to predictive models. Causal modelling, acknowledged a priority at the international level (DARPA 2015), opens principled and sound ways to deal with the unbounded expectations / irrational exuberance about Big Data. Furthermore, causality offers an operational framework to better handle transfer learning, semi-supervised learning and missing data.

Mainstream approaches to causality involve restrictive assumptions (no confounders; no causal cycles) with severe scalability limitations [83]. The international challenges proposed by I. Guyon in the last 3 years (See book in preparation) opened brand new research directions, based on learned causation models [78], [75]. The validation of causal graphs still is an open problem in the general case (multiple hypothesis testing issues, heterogeneous variables, temporal dimensions).

TAU is one of the first teams worldwide with expertise in this domain, collaborating with Max Planck Institute (B. Schölkopf), FORTH (I. Tsamardinos) and Facebook Research (D. Lopez Paz). Among the applications calling for causal models are Energy Management (RTE use-cases include failures of equipment and catastrophic cascades of failures; Inria post-doc work of Berna Batu, see Section 4.1) and computational social sciences (with impact on strategic societal issues, see Section 4.2).

3.2. Deep learning

Deep Learning is at the root of quite a few breakthroughs in machine learning and sequential decision making, albeit requiring gigantic resources [72]. Some reasons for these performance jumps are clear (more data, more computational power, more complex search space). Still, the nature of the dynamical system made of training a deep NN yet remains an open question, at the crossroad of information geometry and non-convex optimization. A related open question concerns the neural architecture design. Deep Learning recent developments regarding generative adversarial networks [68] and domain adaptation [67] are relevant to optimal design applications.

The challenges addressed by TAU range from theoretical ML issues (characterization of learnable problems w.r.t the ratio of the data size/neural architecture size) to functional issues (how to encode information invariance and deal with higher order logic beyond convolutional architectures) to societal issues (how to open the black-box of a deep NN and ensure the fairness of the process).

The TAU team has a unique international expertise in three aspects relevant to deep learning, respectively regarding Riemannian geometry [76], [77] (in order to efficiently navigate in the search manifold), statistical physics [66] (to apprehend the learnability region as the architecture size goes to infinity with the data size), and Genetic Programming [57] and neuro-evolution (that provide original avenues for DNN architecture learning). Related industrial contracts involve ADAMME (FUI 2016) and RTE (Energy Management).

3.3. Optimization and Meta-optimization

TAO, with a first-rank expertise worldwide in stochastic black-box optimization, has now been splitted into the new team RANDOPT, and the present team TAU. While RANDOPT further investigates single- and multi-objective continuous stochastic optimization, TAU continues to focus on the fruitful hybridization of ML and stochastic optimization, with the dual persepectives of using ML for a better informed Optimization, and using Optimization to improve ML performances.

One long-term research perspective in the former context is to apprehend the black-box optimization process (BBO) as a sequential optimal decision process, along the lines of the learning to learn framework [58]. An effective policy (in expectation) can be trained on a representative set of benchmark problems, noting that comparison-based BBO methods offer good generalization properties thanks to their invariances properties, opening the road to Riemanian geometric approaches [11]. Another research perspective concerns interactive optimization, where the initially unknown optimization objective is gradually estimated based on the feedback of the human in the loop, and tackled [62], [56]. But this requires making a trade-off between the optimization search space (rich enough to contain good solutions) and the preference search space (simple enough to support effective preference learning with a limited number of queries).

On the other hand, the meta-optimization problem, concerned with selecting a nearly optimal algorithm and its hyper-parameters depending on the problem instance at hand, has been identified a key issue in both ML [61], [59], and Optimization [71]. This issue becomes a bottleneck for the tranfer to industry, due to the acknowledged shortage of data scientists and the increasing complexity of ML/Optimization toolboxes. The *a priori* algorithm selection and calibration in ML is hindered by the lack of appropriate meta-features to describe a problem instance [10], and the state of the art thus relies on Bayesian optimization, alternatively building a surrogate model of algorithm performances on the instance at hand [82]. The search for meta-features can be revisited, exploiting latent representations derived from Collaborative Filtering [10] and Domain Adaptation approaches based on adversarial networks [68], [67].
Note that Isabelle Guyon was the main organizer of the AutoML challenge, whose purpose was to come up with automatic use of ML methods.

3.4. Big Data-Driven Design

Big data-driven modelling/assimilation/simulation/design (BD3) is concerned with the calibration and extension of first principle-based models and equations using data (aka data assimilation), and using such models for optimal design. BD3 can significantly decrease time-to-design, through fast interactions between the modelling, predicting, optimizing, controlling and designing stages, sharing their advances (in particular, coupling first principles and data [63], or repairing/extending closed-form models). Besides the predictive modelling aspects, TAU more specifically investigates the generative and adversarial modelling aspects [68], aimed at data assimilation from biased data.

A first challenge is to find an operational umbrella to handle noisy, sparse, unstructured, missing data, possibly issued from different distributions (e.g. simulated vs real-world data). Collaborative filtering, deep learning, and their hybrids can be used to forge scalable unified intermediate representations, with applications in energy and computational social sciences (involving times series, documents, and/or graphs). Related issues regard the interpretation of such latent representations and the decisions based thereupon. Another challenge is to deliver guarantees for the data-driven models and designs. As more intelligence is put in the modelling, more intelligence must be put in the validation, as reminded by Leon Bottou. Along this way, generative models will be used to support the design of "what if" scenarios, to enhance anomaly detection and monitoring via refined likelihood criteria.

Several recent, on-going, or submitted projects witness the links of TAU members with experts from application domains: in High Energy Physics (LAL, CERN), in space weather (CWI), in anomaly detection (Thalès ThereSIS), and, within the ADAMME project (FUI 2016), in automatic image labelling (Armadillo), and in yield management (VoyagesSncf.com Technologies).

3.5. Transverse Activity: Organisation of Challenges

Challenges have been an important drive for Machine Learning research for many years, and TAO members have played important roles in the organization of many such challenges: Michèle Sebag was head of the challenge programme in the Pascal European Network of Excellence (2005-2013); Isabelle Guyon, as mentioned, was the PI of many challenges ranging from causation challenges [69], to AutoML [70]. The Higgs challenge [55], most attended ever Kaggle challenge, was jointly organized by TAO (C. Germain), LAL-IN2P3 (B. Kegl) and I. Guyon (not yet at TAO), in collaboration with CERN and Imperial College. The challenge activity continue d within TAU, in relation with fundamental and applied issues.

TAU is particularly implicated with the ChaLearn Looking At People (LAP) challenge series in computer vision, in collaboration with the University of Barcelona [46]. Notably in 2017, TAU co-organized several international LaP challenges:

- ChaLearn Looking at People (LAP) - Job Candidate Screening Coopetition [21]. In conjunction with IJCNN 2017 workshop on explainability in machine learning.
- ChaLearn LAP Real Versus Fake Expressed Emotion Challenge (ICCV 2017) [30].
- ChaLearn LAP Large-scale Continuous Gesture Recognition Challenge (ICCV 2017) [30].
- ChaLearn LAP Large-scale Isolated Gesture Recognition Challenge (ICCV 2017) [30].

TAU was also implicated in organizing a follow up of the AutoML challenge for the PAKDD conference. TAU also co-organized local events (hackathons), as "rehearsals" of international competitions in preparation:

- Spatio-temporal time series challenges for the European See.4C challenge about Energy Management (Paris, 14/2/2017, and Toulon, 22/4/2017). Book with Sprimger in preparation.
- Track ML: tracking particles in high energy physics (Orsay, 21/3/2017) [16].

The Codalab challenge platform, originally designed within Microsoft Research with Isabelle Guyon as one of the PIs, has now been migrated to U. Paris-Sud. It is an open source project. Part of the development is supported by Isabelle Guyon's Paris-Saclay chair (co-funded by Inria). Codalab's user base has been steadily growing. At the end of 2017, we now have over 10'000 users who have entered more than 480 challenges (145 of which are public).

This year, there was a major upgrade of Codalab, featuring:

- A step-by-step Wizard to guide beginner challenge organizers through the process of organizing challenges. This Wizard facilitates the work of students learning to organize challenges.
- Use of dockers and queues, allowing challenge participants to easily use their own computer resources in the backend to support challenges with code submissions.
- A modular competition logic, which will enable supporting new types of challenges such as reinforcement learning competitions.

4. Application Domains

4.1. Energy Management

Energy management has been one of our priority application fields since 2012, under the lead of Olivier Teytaud. The first works were concerned with sequential decision making, and were based on TAO experience in games, in particular GO, starting with the Associated Team (EA) with Tainan (Taiwan) and the Inria ILAB Metis, in collaborations with SME Artelys. This collaboration continued to be very fruitful, with the ADEME BIA project POST (2014-2017), about long-term investments in power systems, and the ADEME BIA NEXT, that started in April 2017 for 4 years, about the optimization of local grids (at the city or region level). Another line of research is addressed in collaboration with RTE, the company that manages the global French electric network, through Benjamin Donnot's CIFRE PhD.

The collaboration with Artelys had moved from sequential decision making in the Metis ILAB to reinforcement learning, and the design of the Direct Policy Search approach to handle non-anticipativity, in the POST project. Currently, the NEXT project is concerned with the optimization of local networks to meet customer demand, and hightlights the need for an accurate, robust, and fast simulator (Big Data), and some efficient modeling of the demand (Small Data). This is the topic of Victor Berger's PhD (started Oct. 2017). Another issue is directly related to the network optimization - and the optimal setting (possibly online) of graph optimization algorithms, which this is the topic of Herilalaina Rakotoarison's PhD, started Nov. 2017.

The on-going collaboration with RTE is about learning the parries in reaction to network or demand changes to enforce the "n-1" security constraint: at any time, the failure of any of the 30000 links in the network should preserve the security constraints. Logs of network operations over many years are available, but without any "parry" label. This can be achieved by simulating what would have happened without that particular operations regarding the n-1 constraint. The available network simulator is far too slow and sensitive to noise to be useful here. Modeling the network using Deep Networks is straightforward, for a given topology, though computationally costly. The challenge is to take into account the tolopology so that the n-1 constraint can be quickly checked with a single network. The first results on a small grid (118 nodes) outperform the classical DC approximation while providing a significant speedup in calculations [42]. Further works include scaling up, and incorporating all the intricacies of real data.

Several other energy-related works have been, or will be addressed [20], including the organization of a large scale challenge funded by the EU, which was endowed with 2 million euros in prizes (Isabelle Guyon co-organizer), in the context of the EU project See.4C. The participants are asked to predict the power flow on the entire French territory over several years. This challenge will be followed by a challenge in reinforcement learning (RL), in the context of Lisheng Sun's PhD thesis (started Oct. 2016), who is now working on the problem of RL and Automatic Machine Learning (reducing to the largest possible extend human intervention in reinforcement learning). Another direction being explored is the use of causal models to improve explainability of predictive models in decision support systems (Inria-funded post-doc Berna Batu). This should allow us making more intelligible suggestions of corrective actions of operators to bring network operations back to safety when incidents or stress occur.

4.2. Computational Social Sciences: Toward AI Fairness

Several TAU projects are related to computational social and economic sciences. This activity is at the core of the French DataIA *Institut de Convergence*, (head Nozha Boujemaa), gathering 19 partners in the Paris-Saclay area to explore the scientific and ethical impacts of data science and artificial intelligence on the academic, industrial and societal sectors.

Many projects in the domain are related to Causal Modelling (see Section 7.1.1). Some are internal to our team; others involve collaborations with external partners, with a transfer dimension. Others are closely related to some Software platform and are desribed in the corresponding Sections (io.datascience, Section 6.1 and Catolabe, Section 6.3).

- AmiQap (Philippe Caillou, Isabelle Guyon, Michèle Sebag, Paola Tubaro, started 2015). The multivariate analysis of state questionaire data relative to the quality of life at work, in relation with the socio-economical indicators of firms, aims at investigating the relationship between quality of life and economic performances (conditionally to the activity sector), in collaboration with the RITM (U. Paris-Sud), SES (IMTelecom) and La Fabrique de l'Industrie, on data gathered by the Ministry of Labour (DARES). AmiQap is a motivating application for the Causal Modelling studies (PhD Divyan Kalainathan; post-doc Olivier Goudet; coll. David Lopez-Paz, Facebook AI Research).
- Collaborative Hiring (Philippe Caillou, Michèle Sebag, started 2014). Thomas Schmitt's PhD, started in 2014, aims at matching job offers and resumes viewed as a collaborative filtering problem. An alternative approach based on Deep Networks has been developped by François Gonard within his IRT SystemX PhD. The study has been conducted in cooperation with the Web hiring agency Qapa and the non-for-profit organization Bernard Gregory.
- U. Paris-Saclay Nutriperso IRS (Philippe Caillou, Flora Jay, Michèle Sebag, Paola Tubaro) aims to uncover the relationships between health, diets and socio-demographic features. The ultimate goal is to provide personalized *acceptable* recommendations toward healthier eating practices. A milestone is to uncover the causal relationships between diet and health (coll. INRA, INSERM, CEA).
- RESTO (Paola Tubaro, Philippe Caillou). A study of transformations brought about by digital platforms and their effects on the restaurants sector, using a mix of methods that includes both agent-based simulations and machine learning, and fieldwork.
- Sharing Networks (Paola Tubaro, started 2016). Mapping the "collaborative economy" of internet platforms through social network data and analysis.
- OPLa - DiPLab (Paola Tubaro). Two related projects investigating the economy of micro-work platforms in France, and how they integrate with the AI industry ecosystem.

Scientific challenges are related to the FAT (Fairness, Accountability and Transparency) criteria: Metric learning, where the distance/topology to be learned must reflect prior knowledge (e.g. ontologies); Interpretation of clusters built from heterogeneous textual and quantitative data, using the learnt metric/distance; Integration of the human-in-the-loop ("dire d'experts"); Assessment of the models w.r.t. their causality (as opposed to their predictive accuracy) in order to support further interventions.

4.3. High Energy Physics (HEP)

The project started in 2015 with the organization of the Higgs boson ML challenge, in collaboration with the Laboratoire de l'Accelerateur Lineaire (LAL) (David Rousseau and Balazs Kègl) and the ATLAS and CMS projects at CERN. These collaborations have been at the forefront of the broadening interaction between Machine Learning and High Energy Physics, with co-organisation of the Weizmann Hammers and Nails 2017 workshops [44], DataScience@HEP at Fermilab and the Connecting The Dots series.

1. **SystML** (Cécile Germain, Isabelle Guyon, Michèle Sebag, Victor Estrade, Arthur Pesah): Experimental data involve two types of uncertainties: statistical uncertainty (due to natural fluctuations), and systematic uncertainty (due to "known unknowns" such as the imprecise characterization of physics parameters). The SystML project aims to deal with experimental uncertainties along three approaches: i) better calibrating simulators; ii) learning post-processors aimed to filter out the system noise; iii) anticipating the impacts of systematic noise (e.g., on statistical tests) and integrating this impact in the decision process.

   V. Estrade's PhD, focusing on the second approach, searches for new data representations insensitive to system-related uncertainty. Taking inspiration from the domain adaptation literature, two strategies have been investigated: i) an agnostric approach based on adversarial supervised learning is used to design an invariant representation (w.r.t. the physics parameters); ii) a prior knowledge-based approach.
2. **TrackML** (Cécile Germain, Isabelle Guyon):
   A Tracking Machine Learning challenge (TrackML) [79], [51] is being set up for 1T 2018. Current methods used employed for tracking particles at the LHC (Large Hadron Collider) at CERN will be soon outdated, due to the improved detector apparatus and the associated combinatorial complexity explosion. The LAL and the TAU team have taken a leading role in stimulating both the the ML and HEP communities to renew the toolkit of physicists in preparation for the advent of the next generation of particle detectors.

   TrackML refers to recognizing trajectories in the 3D images of proton collisions at the Large Hadron Collider (LHC) at CERN. Think of this as the picture of a fireworks: the time information is lost, but all particle trajectories have roughly the same origin and therefore there is a correspondence between arc length and time ordering. Given the coordinates of the impact of particles on detectors (3D points), the problem is to "connect the dots" or rather the points, i.e. return all sets of points belonging to alleged particle trajectories [16]. From the machine learning point of view, beyond simple clustering, the problem can be treated as a latent variable problem, a tracking problem, or a pattern de-noising problem. A very large dataset (100GB) has been built by the Atlas and CMS collaborations specifically for the challenge.

   TrackML will be conducted in 2 phases, the first one favoring innovation over efficiency and the second one aiming at real-time reconstruction. The challenge is supported by Kaggle.

4.4. Autonomous Vehicle

This new application domain builds in fact upon former collaborations of the TAO team with the automotive industry, that created the links with some of the researchers of the R&D departments of Renault (within the Systematic CSDL project and the SystemX ROM project (François Gonard's PhD) and PSA (M. Yagoubi's PhD [84], [85]).

The current work, in collaboration with Renault, is related to the safety of the autonomous vehicle. The validation of the software system is today based on statistics of incidents (failures of some automatized component) assessed from millions of hours of 'driving', either by human drivers in real cars, or by simulations. The work for TAU is related to the set of sample scenarii that are used to compute these statistics. This will require in the first place to identify some latent representation space common to both the actual real-life experiments and the results of the simulation, something that will be achieved using Deep Auto-Encoders of the time series recording the experiments. Two works have started this Fall:

- How to assess the representativity of current set of scenarii, and identify new scenarii to be fed into the simulator to improve the coverage of the scenario space in the common latent representation space, and is the goal of the yet-to-be-signed POC with Renault (Raphaël Jaiswal is working on Renault data since September 2017);
- How to identify original scenarii that lead to failures, an optimization problem in the scenario space. Several criteria for failures will be considered (e.g., getting too close to the preceding car), and the optimization will most likely require building a surrogate model of the simulator for each chosen criterion (and here again Deep Networks are a good candidate), due to its high computing time. This is the topic of Marc Nabhan's CIFRE PhD, started in October 2017 (after a 3 months internship).

4.5. Population Genetics

Work in this application domain started recently, with two main lines of research : dimension reduction of genetic datasets and prediction tasks using genetic data (such as the prediction of past human demography).

- Flora Jay collaborated with Kevin Caye and colleagues (TIMC-IMAG, Grenoble) who developed an R package for inferring coefficients of genetic ancestry, using matrix factorization, alternating quadratic programming and projected least squares algorithms [4]. The extension of ancestry inference and visualization methods to temporal data (for paleogenetics applications) remains to be done.
- The demographic history of one or several population (of any organism) can be partially reconstructed using modern or ancient genetic data. A common approach in the population genetics field is to simulate pseudo-datasets for which the demographic parameters are known and summarize them into handcrafted features. These features are then used as a reference panel in an Approximate Bayesian Computation (likelihood-free) framework. Flora Jay has been developping such methods for the application to whole-genome data [14] [60].
- An untackled challenge in the field is to skip the summary step and directly handle raw data of genetic variations. Théophile Sanchez, who did a 6 month internship in TAU, started his PhD in October 2017 and is currently designing deep learning architectures that are suitable for multi-genome data [33]. In particular these networks should be invariant to the permutation of indiviual genomes and flexible to the input size (see Section 7.2.7).

5. Highlights of the Year

5.1. Organisation and Distinctions

- Isabelle Guyon, General Chair, **NIPS 2017** in Los Angleles (8000+ attendees). She also co-organized several workshops (two *See.4C* workshops, *Connecting the dots* at LAL, *AutoML* a ICML, BayLearn, and CiML at NIPS).
- Flora Jay co-organized **JDSE17**, the second edition of the Junior Conference on Data Science and Engineering, Paris-Saclay (September 2017).
- Yann Ollivier coordinated several events in France (workshop, public conferences, initiatives with school teachers, ...) related to **Shannon100**, the celebration of the Claude Shannon's hundredth birthday, a world-wide event. In particular he created a public exhibit that took place from December 2016 to April 2017 in the Musee des Arts et Metiers in Paris, with extremely positive feedback.
- Marc Schoenauer, expert with Cédric Villani for his national mission on the French AI strategy.
- Michèle Sebag, elected at the *Académie Française des Technologies*; ephemeral nominated member of the *Conseil National du Numérique* (Dec. 2017); member of TransAlgo; head of the DataIA Research programme.

- Paola Tubaro organized **RECSNA17**, an international conference on Recent Ethical Challenges in Social Network Analysis with support from Maison des Sciences de l'Homme Paris-Saclay and Institute for Advanced Studies, in partnership with British Sociological Association, Association Française de Sociologie and European Network on Digital Labor.

5.2. Awards and Prizes

- **AS-AC-CMA-ES Winner**, single objective track at **BBComp**, the Black Box Competition for continuous optimization at ACM-GECCO 2017 (July, Berlin). Nacim Belkhir, Johann Dréo, Pierre Savéant and Marc Schoenauer.
- **ASAP** V2 and V3 [23] ranked **first and second** at the **Open Algorithm Selection Challenge 2017** (see the official results – slide 22). François Gonard, Marc Schoenauer, and Michèle Sebag.

BEST PAPERS AWARDS :

[28] **GSI 2017 - 3rd conference on Geometric Science of Information**. Y. OLLIVIER, G. MARCEAU-CARON.

6. New Software and Platforms

6.1. io.datascience

Input Output Data Science

KEYWORDS: Open data - Semantic Web - FAIR (Findable, Accessible, Interoperable, and Reusable)

FUNCTIONAL DESCRIPTION: io.datascience (Input Output Data Science) is the instance of the Linked Wiki platform developed specifically in Paris-Saclay University as part of its Center for Data Science.

The goal of io.datascience: to facilitate the sharing and use of scientific data. The technological concept of io.datascience: the exploitation of semantic web advances, and in particular wiki technologies.

One of the grand challenges of data-intensive science is to facilitate knowledge discovery by assisting humans and machines in their discovery of, access to, integration and analysis of, task-appropriate scientific data and their associated algorithms and workflows. The guiding principles for this challenge have been defined: Data should become FAIR (Findable, Accessible, Interoperable, and Reusable) (Wilkinson, M., and The FAIR Guiding Principles for Scientific Data Management and Stewardship, Nature Scientific Data 2016)

io.datascience is both a data sharing platform and a framework for further development. It realizes a practical implementation of FAIR principles through a user-centric approach. • Share: Software users can declare the sources of the data they use as well as their query requests. • Discover: Using a form, users can link their data sources to each other. The repository used is that of Wikidata. The user can then retrieve his data sources and example queries through a search interface or directly through Google and Wikipedia. • Reuse: data is identified and qualified, a simple interface allows the user to provide the desired level of description for the data they refer to, as well as examples of use. • Analyze: io.datascience will soon be proposing the creation of RDF databases on the cloud on the cloud of Paris Sud University.

- Partners: Border Cloud - Paris Saclay Center for Data Science - Université Paris-Sud
- Contact: Cécile Germain-Renaud
- Publications: Data acquisition for analytical platforms: Automating scientific workflows and building an open database platform for chemical anlysis metadata - A platform for scientific data sharing - TFT, Tests For Triplestores - Une autocomplétion générique de SPARQL dans un contexte multi-services - Certifying the interoperability of RDF database systems - Transforming Wikipedia into an Ontology-based Information Retrieval Search Engine for Local Experts using a Third-Party Taxonomy - The Grid Observatory 3.0 - Towards reproducible research and open collaborations using semantic technologies
- URL: https://io.datascience-paris-saclay.fr/

6.2. Codalab

KEYWORDS: Benchmarking - Competition
FUNCTIONAL DESCRIPTION: Challenges in machine learning and data science are competitions running over several weeks or months to resolve problems using provided datasets or simulated environments. Challenges can be thought of as crowdsourcing, benchmarking, and communication tools. They have been used for decades to test and compare competing solutions in machine learning in a fair and controlled way, to eliminate "inventor-evaluator" bias, and to stimulate the scientific community while promoting reproducible science. See our slide presentation.

As of december 2017 there are 145 public competitions on Codalab and over 10000 users. Some of the areas in which Codalab is used include Computer vision and medical image analysis, natural language processing, time series prediction, causality, and automatic machine learning. Codalab was selected for the million Euro challenge See.4C that was awarded a H2020 EU grant for its organization.

TAU is going to continue expanding Codalab to accommodate new needs. One of our current focus is to support use of challenges for teaching (i.e. include a grading system as part of Codalab) and support for hooking up data simulation engines in the backend of Codalab to enable Reinforcement Learning challenges and simulate interactions of machines with an environment. For the third year, we are using Codalab for student projects. M2 AIC students create mini data science challenges in teams of 6 students. L2 math and informatics students then solve them as part of their mini projects. We are collaborating with RPI (New York, USA) to use this platform as part of a curriculum of medical students. Our PhD. students are involved in co-organizing challenges to expose the research community at large with the topic of their PhD. This helps them formalizing a task with rigor and allows them to disseminate their research.

- Partner: Microsoft
- Contact: Isabelle Guyon
- URL: http://competitions.codalab.org

6.3. Cartolabe

KEYWORD: Information visualization
FUNCTIONAL DESCRIPTION: The goal of Cartolabe is to build a visual map representing the scientific activity of an institution/university/domain from published articles and reports. Using the HAL Database and building upon the AnHALytics processing chain, Cartolabe provides the user with a map of the thematics, authors and articles and their dynamics along time. ML techniques are used for dimensionality reduction, cluster and topics identification, visualisation techniques are used for a scalable 2D representation of the results.
NEWS OF THE YEAR: Improvement of the graphical interface

- Contact: Philippe Caillou
- URL: http://cartolabe.lri.fr/

7. New Results

7.1. Causality, Explainability, and Reliability

As said, the fairness, accountability and transparency of AI/ML need be assessed, measured and enforced to address the ethical impacts of data science on industry and society. TAU has started working toward improving the confidence in ML algorithms through three research directions.

7.1.1. Causality

Links between quality of life at work and company performance Within the Amiqap project, a new approach to functional causal modeling from observational data called *Causal Generative Neural Networks* (CGNN) has been developed [45]. CGNN learns a generative model of the joint distribution of the observed variables, by minimizing the Maximum Mean Discrepancy between generated and observed data. An approximate learning criterion scales the computational cost of the approach to linear complexity in the number of observations. CGNN extensions, motivated by the redundancy of real-world variables, are under-going to achieve a causal model of the corporate- and human resource-related variables at the firm and economic sector levels.

Generating Medical Data This project, in collaboration with RPI (New York), aims to provide medical students with case studies, generated using CGNN . and fully preserving their confidentiality. We are exploring the benefits of using data generated by CGNNs in replacement for real data. Such data will preserve the structure of the original data, but the patient records will not represent real patients.

Missing Data Missing and corrupted data is a pervasive problem in data modeling. Our interest in this problem stems from 2 applications: epidemiology (in collaboration with Alain-Jacques Valleron, INSERM, and RPI New York) and computer vision (in collaboration with Aix-Marseille University and University of Barcelona). As it turns out, missing data is a causality problem [80]. In a paper under review, we outline the danger of imputing values in risk factor analysis in the presence of missing data. We are also preparing a challenge on the problem of "inpainting" to restore images with occlusions and to eliminate captions in movies.

Power Networks Berna Batu (post-doc Inria) explores causal modeling in time series to explain cascades of events. Other applications (e.g., in epidemiology) may develop from this approach.

7.1.2. Explainability

Explainable Machine Learning for Video Interviews [21]. The challenge consisted in analyzing 15s videos, (human) annotated with the Big Five persinality traits (Openness to experience, Conscientiousness, Extroversion, Agreableness, and Neurotism – sometines referred to as OCEAN features). Human annotators also voted whether a given candidate should be invited for an interview. As organizers we provided a strong baseline system, which was based on deep learning methods having won part challenges. Onty the winners outperformed quantitatively the basline method.

The winner of the prediction challenge (BU-NKU) performed a very sophisticated analysis, combining face analysis (from the entire video) and scene analysis (from the first image), both analyses contributing to the final decision. Face analysis extracted spatio-temporal featured from a pre-trained convolutional neural network (CNN) and using Gabor filters. Scene analysis features were also extracted with a pre-trained CNN. Acoustic features were extracted with the OpenSMILE tool. From the feature set, the personality traits are predicted with kernel ridge regression and from there on, the "invite for interview" is predicted using Random Forests.

For the explainability challenge, the BU-NKU team performed final predictions with a classifier based on binarized predicted OCEAN scores mapped to the binarized ground truth using a decision tree, a self-explanatory model that can be converted into an explicit recommender algorithm, using the trace of each decision from the root of the tree to the leaf. The verbal explanations are finally accompanied with the aligned image from the first face-detected frame and the bar graphs of corresponding mean normalized scores. Trained on the predicted OCEAN dimensions, this gave over 90% classification accuracy.

Note that another team (TUD), who did not enter the quantitative competition, nevertheless won forst place ex-aequo with the BU-NKU team on the explainability challenge. Interestingly, they added facial features (using OpenFace) and text features (using published "Readability" features) in an effort to capture level of education from the sophistication of language, which was not captured by personality traits. They then used PCA to reduce dimention, and the coefficients of a linear regression model, fed back into the PCA model to generate explanations.

Skin image classification Also, the on-going collaboration with Roman Hossein Khonsari, surgeon at Necker hospital, is continuing, on the topic of skin disease image classification, with the goal of explaining how the trained neural networks produce their predictions, in order to be trusted by users. For this, we analyse the features that are learned, and show which ones are found in each image example.

7.1.3. Model systematic bias and reliability

A related problem is the reliability of models and their robustness to bias. We initiated research on this topic in the context of eliminating bias of Hight Energy Physiscs simulators. Discovering new particules relies on making accurate simulations of particle traces in detectors to diagnose collision events in high energy experiments. We are working on data from the ATLAS experiment at CERN, in collaboration with David Rousseau at the Laboratoire de l'Accelerateur Lineaire (LAL). We produced two preliminary studies on this topic: Adversarial learning to eliminate systematic errors: a case study in High Energy Physics [32] and Robust deep learning: A case study [43].

This line of research will extend to the calibration of other simulators, particularly energy transport and distribution simulators and medical data simulators, which we are working on in the context of other projetcs.

Beyond the calibration of simulators, we are also interested in using such approaches to forter fairness and de-bias data. For instance, in the "personality trait" data mentionned in the previous section, our analysis shows that labelers are biased favorably towards females (vs. males) and unfavorably towards African-American (vs. Caucasian or Asian).

7.2. Deep Learning and Information Theory

7.2.1. Convergence proofs for recurrent networks

Pierre-Yves Massé, in his PhD, defended Dec.2017 under the supervision of Yann Ollivier [3], obtained the very first rigorous results of convergence for online training of recurrent neural networks, by viewing them from the viewpoint of dynamical systems.

7.2.2. Fast algorithms for recurrent networks

Corentin Tallec (in his on-going PhD) and Yann Ollivier produced a new, faster algorithm for online training of recurrent networks, UORO, which is guaranteed to converge locally, and requires only linear time [49].

7.2.3. An explanation for LSTMs

The LSTM structure is currently the most popular recurrent network architecture. However, it is quite complex and very much ad hoc. Corentin Tallec (in his on-going PhD) and Yann Ollivier derived this architecture from first principles in a very simple axiomatic setting, simply by requiring that the model is invariant to arbitrary time deformations (such as accelerations, decelerations) in the data.

7.2.4. Bayesian neural networks

The Bayesian approach to neural networks makes several suggestions. First, it suggests to artificially add a very specific amount of noise during training, as a protection against overfit. This has to be done carefully (Langevin dynamics) in relation with the Fisher information metric. Gaetan Marceau-Caron and Yann Ollivier demonstrated that this approach can be applied efficiently for neural networks [28] (Best paper award at GSI17).

Second, a Bayesian viewpoint can help select the right size for each layer in a neural network. A comparison to a theoretical model of an infinitely large network suggests ways to adapt learning rates and criteria to select or deselect neurons or even layers (Preliminary results in a preprint by Pierre Wolinski (PhD), Yann Ollivier and Guillaume Charpiat, in preparation.)

7.2.5. Kalman filtering and information geometry

Filtering and optimization have been brought much closer by the following result [48]: the natural gradient in optimization is mathematically fully identical to the Kalman filter, for all probabilistic (machine learning) models. Even though both methods had been known for decades and were an important reference in their respective fields, they had not been brought together. The result extends to the non-iid setting (recurrent neural networks).

7.2.6. Computer vision

The activity of computer vision is run jointly with the program of Looking at People (LaP) challenges [46]. We edited a book in Springer, which is a collection of tutorials and papers on gesture recognition [54], to which we contributed a survey chapter on deep-learning methods [34] a shorter version of which was published at the FG conference [17].

Several papers were published this year analyzing past LAP challenges. The “first impressions” challenge aimed at detecting personality traits from a few seconds of video. In [8], we demonstrate how deep residual networks attain state-of-the art performance on that task and lend themselves well to identifying which parts of the image is responsible for the final decision (interpretability). We also analyzed last years’ challenge on apparent age estimation from in still images and proposed improvements with deep residual networks [15]. A similar methodology based on deep-residual networks was applied to apparent personality trait analysis [24], [8].

7.2.7. Flexible deep learning architectures suitable to genetic data

Genetic data is usually given in the form of matrices, one dimension standing for the different individuals studied and the other dimension standing for the DNA sites. These dimensions vary, depending on the indivual sample size and on the DNA sequence length. On the other side, standard deep learning architectures require data of fixed size. We consequently search for suitable, flexible architectures, with as an application the prediction of the demographic history of a population given its genetic data (changes in the number of individuals through time). Théophile Sanchez, now PhD student, presented his work at the Junior Conference on Data Science and Engineering at Paris-Saclay [33]. To our knowledge this is the first attempt in the population genetics field to learn automatically from the raw data.

7.2.8. Image segmentation and classification

Emmanuel Maggiori, PhD student in the Titane team, Inria Sophia-Antipolis, mainly supervised by Yuliya Tarabalka, and co-supervised by Guillaume Charpiat, defended his PhD thesis [73], on the topic of remote sensing image segmentation with neural networks. This year, an architecture for proposed to be able to deal with high resolution images; a benchmark was built and made public (as there is lack of those in the remote sensing community); and the output of segmentation predictions was turned into a vectorial representation by suitable automatic polygonization [9], [25], [26].

Through a collaboration with the company Armadillo within the ADAMme project, we have also worked on image classification with multiple tags. The database consists of 40 millions images, with thousands of different possible tags (each image is on averaged associated with 10 tags). We started from a ResNet pre-trained network and adapted it to our task. A demonstration of our results was performed at the annual review meeting of the project.

7.2.9. Non-rigid image alignment

Automatic image alignment was also studied. In remote sensing, the task consists in aligning satellite or aerial images with ground truth data such as OpenStreetMap’s cadastrial maps. This task is crucial in that such ground truth data is actually never well registered but is spatially deformed, preventing any further use by machine learning tools. Based on the analysis of multiple scale classical frameworks, a deep learning architecture was proposed to perform this task. This work is currently under submission to CVPR. On a related topic, in a collaboration with the start-up company Therapixel, we have been studying the registration of 3D medical images, but without any ground truth or template.

7.2.10. Video analysis

Time coherency is usually poorly handled in video analysis with neural networks. We have studied, on 3 different applications, different ways to take it better into account. First, in a collaboration with the Vision Institute, we studied different ways of incorporating neural networks in reinforcement learning approaches for the tracking of microbes with a motorized microscope. Second, in a collaboration with the SATIE team, we worked on the incorporation of optical flow for crowd density estimation, and, finally, in a collaboration with the Parietal team, we study how to link brain fMRI signals to the videos people are watching.

7.3. Algorithm Selection and Configuration

Automatic algorithm selection and configuration (hyper-parameter selection) depending on the problem instance at hand is a pervasive research topic in TAO, for both fundamental and practical reasons: in order to automatically deliver a peak performance on (nearly) every new problem instance, and to understand the specifics of a problem instance and the algorithm skills w.r.t. these specifics.

7.3.1. Algorithm recommendation

A collaborative filtering approach called Alors (Algorithm Recommender System) has been proposed to achieve algorithm selection [10], considering after [81] that a problem instance "likes better" algorithms with good performances on this instance. Alors, tackling a cold-start recommendation problem, enables to independently assess the quality of the benchmark data (representativity of the problem instances w.r.t. the algorithm portfolio) and the quality of the meta-features used to describe the problem instances. Experiments on SAT, CSP and ML benchmarks yield state-of-art performances in the former two domains; these good results contrast with the poor results obtained on the ML domain, blamed on the comparatively poor quality of the ML meta-features.

7.3.2. AutoML and AutoDL

Isabelle Guyon has organized the AutoML challenge (paper in preparation), proposing a series of algorithm selection and configuration problems of increasing difficulty. Following this successful challenge, a new challenge will be organized in collaboration with Google Zurich, specifically targeting the selection of deep network architectures (AutoDL: Automatic Deep learning) in five domains: Image; Video; Audio; Text; Customer demographic descriptors.

The expected result of the challenge is to alleviate the burden on data scientists to design a good architecture ("black art"), and to enforce the reproducibility of the results. In particular, this challenge will encourage advances regarding a few key research questions:

- How to make optimization algorithms more efficient without introducing more tunable parameters?
- How to efficiently automate the tuning of many hyper-parameters?
- How to automatically design or optimize a network architecture for a particular problem?
- How to further automate the learning process by directly learning how to learn?

7.3.3. Per Instance Algorithm Configuration for Continuous Optimization

Nacim Belkhir's PhD thesis (defended on Nov. 30., 2017) was centered on PIAC (Per Instance Algorithm Configuration) in the context of continuous optimization. After a detailed study of features that had been proposed in the litterature, he studied the dependency of the PIAC results on the size of the sample on which they are computed. The rationale is that you must take into account the number of function evaluations that are used to compute the features when addressing a new target instance. He demonstrated that PIAC based on very small sample sets (down to 50 times the dimention) can nevertheless help improving the overall results of the optimization procedure [18], in particular by winning the single-objective track of the GECCO 2017 Black Box Competition.

7.3.4. Feature-based Algorithm Selection in Combinatorial Optimization

In the first part of his PhD (to be defended in Feb. 2018, see also Section 4.2), François Gonard designed ASAP, an Algorithm Selection algorithm that combines a global pre-scheduler and a per instance algorithm selector, to take advantage of the diversity of the problem instances on one hand and of the algorithms on the other hand. ASAP participated to two competitions: the 2016 ICON challenge [35], in which it obtained a *Special Mention* for its originality (and obtained excellent results on half of the problems); the 2017 OASC challenge where two versions of ASAP obtained the first overall best performances [23].

7.3.5. Deep Learning calibration

In a starting collaboration with Olivier Teytaud (who left TAO for Google Zurich in 2016), we proposed [40] an online scheme for Deep Learning hyper-parameter tuning that detects and early-stops unpromising runs using extrapolation of learning curves [64], taking advantage of the parallelism, and offering optimality guarantees within the multiple hypothesis testing framework.

7.3.6. Learning Rate Adaptation in Stochastic Gradient Descent

Based on an analogy with CMA-ES step-size adaptation (comparison with random walks), an original mechanism was proposed for adapting the learning rate of the stochastic gradient descent [52]. As increasing the learning rate can increase the number of catastrophic events (exploding gradients or loss values), a change detection test is used to detect such events and backtrack to safe regions. First experiments on small size problems (MNIST and CIFAR10) validate the approach. Interestingly, the same mechanism can be applied to the Adam optimizer and also improves on its basic version.

7.3.7. Domain Adaptation

The subject of V. Estrade's PhD is to advance domain adaptation methods in the specific context of uncertainty quantification and calibration in High Energy Physics analysis. The problem consists of learning a representation that is insensitive to perturbations induced by nuisance parameters. The need for the adversarial techniques, assuming a completely knowledge-free approach, has been questioned. Our results [32], [43] contrast the superior performance of incorporating a priori knowledge (Tangent Propagation approach) on a well separated classes problem (MNIST data) with a real case setting in HEP.

7.4. Generative Models and Data-driven Design

Learning generative models from observational data faces two critical issues: model selection (defining a loss criterion well suited to the considered distribution space) and tractable optimization.

7.4.1. A Statistical Physics Perspective

Restricted Boltzmann machines (RBM) define generative models, and advanced mean field methods of statistical physics can be leveraged to analyze the learning dynamics. Giancarlo Fissore's Master thesis (now in PhD), co-supervised by Aurélien Decelle and Cyril Furtlehner, has characterized the information content of an RBM from its spectral properties and derived a phenomenological equation of the learning process by means of the spectral dynamics of the weight matrix [5]. The learning dynamics has been analyzed in both linear and non-linear regimes, investigating the impact of the input data.

Secondly [37], the weight matrix ensemble which results from this spectral representation is used to analyze the thermodynamical properties of RBMs in terms of a phase diagram. The conditions for the RBM training, interpreted as a so-called ferromagnetic compositional phase, are given. Ferromagnetic order parameters are identified in the aforementionned phenomenological equation; a closed-form is obtained through explicit integration in simple cases, yielding a behavior of the learning spectral dynamics that matches the actual dynamics of standard RBM training (e.g. using contrastive divergence). After this model, a repulsive interaction takes place among the singular modes of the weight matrix, as some pressure of the lower modes is exerted on higher modes along training. Remarkably, this repulsive interaction is observed in algorithmic experiments for low learning rates.

7.4.2. Functional Brain Dynamics

Generative models have also been used by Aurélien Decelle and Cyril Furtlehner to model the dynamics of the Functional conectome (FCD) in the context of the BRAINTIME exploratory project, along two lines.

On the one hand, Restricted Boltzmann Machines have been used to learn the statistics of the time-varying resting state BOLD activity of 49 human subjects in the age span of 18 to 80 years. RBM models trained on a *per* individual basis show at least two statistically distinct pure states for each subject, between which resting state activity is stochastically wandering. Through mean-field TAP approximations of free energy we have evaluated the energy barrier between these two states *per* individual. Interestingly young and old individuals have different switching statistics: more regular for young subjects *vs* bursty and temporally irregular for elderly subjects. Furthermore, the switching probability is correlated with the energy difference between the two pure RBM states, opening the way to a personalized "landscape" analysis of the resting state FCD.
On the other hand, extremely sparse precision matrices describing the co-activation statistics of different brain regions during resting state based on BOLD time-series have been derived using sparse Gaussian copula models. Such extremely sparse models support direct inter-subject comparisons, in contrast with usually dense FC descriptions. A further step is to characterize the brain activity dynamics, e.g. through considering multi-temporal slice models.

7.4.3. Power systems Design and Optimization

Last work within the POST project, Vincent Berthier's PhD [2] addressed issues in global continuous optimization, and proposed to use unit commitment problems to go beyond classical benchmarks of analytical functions.

Benjamin Donnot (RTE Cifre PhD, now under Isabelle Guyon's supervision), successfully started to disseminate his work in the power system community [20]. His main results regard the design of an original alternative to the one-hot encoding for the topology of the French power grid, termed *Guided Dropout*. Taking advantate of the high redundancy of network connections, the idea is to learn a random mapping between all possible "n-1" topologies and the connections of the neurons [65], [42].

7.4.4. Multi-Objective Optimization

Dynamic Objectives: Within the E-Lucid project, coordinated by Thales Communications & Security, the on-going work about anomaly detection in network flow [74] led to an original approach to many-objective problem, where the objectives are gradually introduced, preventing the population to be abruptly driven toward satisficing only the easy objectives at the beginning of the evolution [27] (runner-up for the Best Paper Award of the Evolutionary Multi-Objective track at GECCO 2017).
Dynamic Fitness Cases: In [22], we propose to gradually introduce the fitness cases in the case of symbolic regression with Genetic Programming, so as to guide the search more smoothly. Experimental results demonstrate a better success rate in the case of both static and dynamic problems.

7.4.5. Space Weather Forecasting

In the context of the MDG-TAU joint team project, focusing on space weather forecasting, Mhamed Hajaiej's Master thesis (under Aurélien Decelle, Cyril Furtlehner and Michèle Sebag's supervision) has tackled the prediction of magnetic storms from solar magnetograms, more specifically considering the representation of solar magnetograms based on auto-encoders. Besides finding a well-suited NN architecture, the difficulty was to find a loss function well suited to the data distribution. A next step (Mandar Chandorkar's PhD at CWI under Enrico Camporeale supervision) is to estimate from the solar images the speed of the solar wind, and the time needed for solar storms to reach the first Lagrange point; this estimation is meant to build a well-defined supervised learning problem, associating a solar image at t to its effect measured at $t + \delta$ on the first Lagrange point.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- **Thales Research & Technology** 2014-2017 (30 kEuros), related to Nacim Belkhir's CIFRE PhD
  Coordinator: Marc Schoenauer
  Participants: Johann Dréo, Pierre Savéant, Nacim Belkhir
- **ESA Tender** 2016-2017 (52 kEuros)
  Coordinator: Bart Boonacker (TNO)
  Participant: Marc Schoenauer, Dejan Tusar
- **Réseau Transport d'Electricité** 2015-2018 (72 kEuros), related to Benjamin Donnot's CIFRE PhD
  Coordinator: Olivier Teytaud (until May 2016), now Isabelle Guyon
  Participants: Benjamin Donnot, Antoine Marot, Marc Schoenauer
- **Therapixel** 2017 (6 mois, 3 kEuros), on the topic of 3D medical image non-rigid registration with neural networks
  Coordinators: Guillaume Charpiat, Olivier Clatz
  Participant: Priyanka Mandikal (master internship)
- **Myndblue**, 2017-2018 (1 an, 50kEuros) related to consulting activities with DMH (Digital for Mental Health).
  Coordinator: Aurélien Decelle
  Participants: all TAU members
- **La Fabrique de l'Industrie** 2017-2018 (1 an, 30kEuros) A COMPLETER (Michéle ?)
- **Renault (POC)** 2017-2018 (125 kEuros), *Clusterisation et optimisation de scenarii pour la validation des véhicules autonomes*
  Coordinator: Marc Schoenauer and Philippe Reynaud (Renault)
  Participants: Guillaume Charpiat, Raphaël Jaiswal (engineer), Marc Schoenauer
- **Renault (CIFRE)** 2017-2020 (45 kEuros), related to Marc Nabhan's CIFRE PhD *Sûreté de fonctionnement d'un véhicule autonome - évaluation des fausses détections au travers d'un profil de mission réduit*
  Coordinator: Hiba Hage and Yves Tourbier (Renault)
  Participants: Guillaume Charpiat, Marc Nabhan (PhD), Marc Schoenauer
- **RESTO** 2017 (14k Euros), *REseaux et Simulations : usages Technologiques et Opinions multiples sur les plateformes numériques dans les marchés de la restauration*, funded by Mission Interdisciplinarité of CNRS. Supported the internship of J. Posada.
  Coordinator: Paola Tubaro
  Participants: Philippe Caillou (with partners at Telecom ParisTech and Université Paris-Dauphine).
- **OPLa** 2017-2018, Organizing Platform Labor (27k euros), funded by Force Ouvrière.
  Coordinator: A.A. Casilli (Telecom ParisTech)
  Participants: Paola Tubaro
- **DiPLab** 2017-2018, Digital Platform Labor (24k euros), funded by MSH Paris-Saclay.
  Coordinators: Paola Tubaro (avec A.A. Casilli, Telecom ParisTech)

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

- **ACTEUR** 2014-2018 (236kEuros). Cognitive agent development for urban simulations,
  Coordinator: P. Taillandier (IDEES, Univ Rouen)
  Participant: Philippe Caillou
- **EPITOME** 2017-2020 (225kEuros). Efficient rePresentatIon TO structure large-scale satellite iM-agEs,
  Coordinator: Yuliya Tarabalka (Titane team, Inria Sophia-Antipolis)
  Participant: Guillaume Charpiat

9.1.2. Others

- **ROM***Model Reduction and Multiphysics Optimization* 2014-2017 (50 Keuros)
  Coordinator: IRT System X
  Participants: Marc Schoenauer, Michèle Sebag, François Gonard (PhD)
- **MAJORE***A Collaborative Filtering Approach to Matching Job Openings and Job Seekers*, 2013-2017 (105 kEuros)
  Thomas Schmitt's PhD (funded by ISN).
  Participants: Philippe Caillou, Michèle Sebag, Thomas Schmitt (PhD)
- **AMIQAP** 2015-2017 (12 months of Postdoctoral fellow). Qualité de vie au travail.
  Project funded by ISN
  Partners: Mines-Telecom SES, RITM (Univ. Paris Sud) and *La Fabrique de l'Industrie*
  Extended for 6 months in 2018 via a donation from *La Fabrique de l'Industrie*
  Participants: Philippe Caillou, Olivier Goudet, Isabelle Guyon, Michèle Sebag, Paola Tubaro, Diviyan Kalainathan (PhD)
- **Nutriperso** 2017-2018, 37 kEuros. Personalized recommendations toward healthier eating practices.
  U. Paris-Saclay IRS (*Initiative de Recherche Stratégique*)
  Partners: INRA (coordinator), INSERM, Agro Paristech, Mines Telecom
  Participants: Philippe Caillou, Flora Jay, Michèle Sebag, Paola Tubaro
- **POST** 2014-2017 (1,220 MEuros, including 500 kEuros for a 'private' cluster). Platform for the optimization and simulation of trans-continental grids
  ADEME (Agence de l'Environnement et de la Maîtrise de l'Energie)
  Coordinator: ARTELYS
  Participants (in 2017, after Olivier Teytaud left): Vincent Berthier (PhD defended in Dec.), Marc Schoenauer
- **E-LUCID** 2014-2017 (194 kEuros)
  Coordinator: Thales Communications & Security S.A.S
  Participants: Marc Schoenauer, Cyril Furtlehner, Luis Marti
- **PIA ADAMME** 2015-2018 (258 kEuros)
  Coordinator: Bull SAS
  Participants (in 2017): Marc Schoenauer, Guillaume Charpiat, Cécile Germain-Renaud, Yasmina Bouzbiba, Etienne Brame
- **CNES contract** 2015-2017 (70 kEuros)
  Coordinator: Manuel Grizonnet (CNES) & Yuliya Tarabalka (Inria Sophia-Antipolis, Titane team)
  Participant: Guillaume Charpiat
- **NEXT** 2017-2021 (675 kEuros). Simulation, calibration, and optimization of regional or urban power grids
  ADEME (Agence de l'Environnement et de la Maîtrise de l'Energie)
  Coordinator: ARTELYS
  Participants Isabelle Guyon, Marc Schoenauer, Michèle Sebag, Victor Berger (PhD), Herilalaina Rakotoarison (PhD), Berna Bakir Batu (Post-doc)

- **BRAINTIME** 2017 (7 kEuros) Défi exploratoire interdisciplinaire de l'appel INFINITE (CNRS) concerning the functional connectome dynamics of the brain.
  Coordinator: Andrea Brovelli (CNRS), Institut de Neurosciences de la Timone (INT)
  Participants Aurélien Decelle, Cyril Furtlehner
- **CDS DeepGenetics** 2017 (6mois, 3k euros), Deep Learning for Population Genetics.
  funded by Center for Data Science
  Coordinators: Flora Jay and Guillaume Charpiat
  Participants: Théophile Sanchez (master internship)

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

See.4C 2016-2017 (2.7 kEuros). SpatiotEmporal ForEcasting: Coopetition to meet Current Cross-modal Challenges
Participants: Isabelle Guyon

9.2.2. Collaborations with Major European Organizations

MLSpaceWeather 2015-2019. Coupling physics-based simulations with Artificial Intelligence.
Coordinator: CWI
Participants: Michèle Sebag, Aurélien Decelle, Cyril Furtlehner, Mhamed Hajaiej

ESA Tender 2015-2017
Coordinator: Bart Boonacker (TNO)
Participant: Marc Schoenauer, Dejan Tusar

9.3. International Initiatives

9.3.1. Inria Associate Teams Not Involved in an Inria International Labs

9.3.1.1. MDG-TAO

Title: Data-driven simulations for Space Weather predictions

International Partner: CWI (Netherlands) – Multiscale Dynamics Group – Enrico Camporeale

Start year: 2017

We propose an innovative approach to Space Weather modeling: the synergetic use of state-of-the-art simulations with Machine Learning and Data Assimilation techniques, in order to adjust for errors due to non-modeled physical processes, and parameter uncertainties. We envision a truly multidisciplinary collaboration between experts in Computational Science and Data assimilation techniques on one side (CWI), and experts in Machine Learning and Data Mining on the other (Inria). Our research objective is to realistically tackle long-term Space Weather forecasting, which would represent a giant leap in the field. This proposal is extremely timely, since the huge amount of (freely available) space missions data has not yet been systematically exploited in the current computational methods for Space Weather. Thus, we believe that this work will result in cutting-edge results and will open further research topics in space Weather and Computational Plasma Physics.

9.3.2. Inria International Partners

9.3.2.1. Declared Inria International Partners

Isabelle Guyon partner of Google Zurich *Preparation of a competition AutoDL: Automatic Deep Learning*.

9.3.2.2. Informal International Partners

Marc Schoenauer partner of the ARC-DP (Australian Research Council Discovery Project) *Bio-inspired computing methods for dynamically changing environments*. Coordinator: University of Adelaide (Frank Neumann), 5 years from Nov. 2015, 400 k$-AUS. Visit to Adelaide: 2 weeks in Feb. 2017. Paper in preparation.

Isabelle Guyon partner of UC Berkeley *Fingerprint verification with deep siamese neural networks using ultratonic sensor data.* Co-advisor of a master student (Baiyu Chen). Partners: Alyosha Efros, Bernhard Boser.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

- **Edgar Galvan Lopez** University College Dublin, April 2015 - April 2017, funded by the ELEVATE Fellowship, the Irish Research Council's Career Development Fellowship co-funded by Marie Curie Actions. Now Lecturer at Maynooth University, Ireland.

9.4.1.1. Internships

- **Tomas Lungenstrass** June 2016 - June 2017, self-funded, collaboration with Inria Chile. Worked on magnetic storm prediction under A. Decelle's, C. Furtlehner's and M. Sebag's supervision.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- Guillaume Charpiat, Workshop Statistics/Learning at Paris-Saclay 2017 and 2018
- Isabelle Guyon, General Chair, NIPS 2017
- Flora Jay, Junior Conference on Data Science and Engineering Paris-Saclay (JDSE 2017)
- Paola Tubaro, Recent Ethical Challenges in Social Network Analysis (RECSNA 2017)

10.1.1.2. Member of the Organizing Committees

- Cecile Germain, co-organizer of DataScience@HEP 2017; Hammers and Nails Weizmann Workshop.
- Isabelle Guyon, co-organizer BayLearn, NIPS workshops Challenges in Machine Learning, AutoML workshop at ICML, LAP challenge workshops (ICCV, ICPR).
- Marc Schoenauer, Steering Committee, Parallel Problem Solving from Nature (PPSN); Steering Committee, Learning and Intelligent OptimizatioN (LION).
- Michele Sebag, President of Steering Committee, Eur. Conf. on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD).

10.1.1.3. Member of Conference Program Committees

All TAO members are members of the Program Committees of the main conferences in the fields of Machine Learning, Evolutionary Computation, and Information Processing.

10.1.1.4. Reviewer

All TAO member review papers for the most prestigious conferences in the fields of Machine Learning and Evolutionary Computation.

10.1.2. Journal

10.1.2.1. Member of the Editorial Boards

- Isabelle Guyon, action editor, *Journal of Machine Learning Research* (JMLR); series editor, *Springer series Challenges in Machine Learning* (CiML).
- Marc Schoenauer, member of Advisory Board, *Evolutionary Computation Journal*, MIT Press, and *Genetic Programming and Evolutionary Machines*, Springer Verlag; action editor, *Journal of Machine Learning Research*(JMLR).
- Michèle Sebag, Editorial Board, Machine Learning, Springer Verlag.
- Paola Tubaro, Associate Editorial Board, *Sociology*, Sage; member of Editorial Board, *Revue Française de Sociologie*, Presses de Sciences Po.

10.1.2.2. Reviewer - Reviewing Activities

All members of the team reviewed numerous articles for the most prestigious journals in the fields of Machine Learning and Evolutionary Computation.

10.1.3. Invited Talks

- Philippe Caillou, 7 march 2017, *Simulation analysis with charts in GAMA*, Gama Training session, TU Delft, Delft.
- Guillaume Charpiat, 9 May 2017, *Introduction to Neural Networks*, Mathematical coffees, Huawei.
- Guillaume Charpiat, 30 November 2017, *Apprentissage profond pour la segmentation d'images satellite haute résolution*, Workshop Deep Learning - Télédétection - Temps, Issy-les-Moulineaux.
- Aurélien Decelle, 24 March 2017, *Ising inverse problem : recovering the topology of the network*, International workshop on numerical methods and simulations for materials design and strongly correlated quantum matters
- Isabelle Guyon, 19 Jan 2017, *Causal graph reconstruction*, ENS Ulm, Paris.
- Flora Jay, 7 April 2017, *Reconstructing past history from whole-genomes: an ABC approach handling recombining data*, European Mathematical Genetics Meeting, Estonia.
- Cecile Germain, 9 May 2017, *Review on Anomaly/Outlier detection*, DataScience@HEP, Fermilab.
- Marc Schoenauer, 23 Feb. 2017, *Adaptation and self-adaptation in Evolutionary Computation and in scientific careers*, School of Computer Science, University of Adelaide; 7 Sep. 2017, *l'Intelligence Artificielle dans le domaine scientifique*, Open Laboratories, IMRA, Sophia Antipolis; 21 Sep. 2017, *Getting hints from random walks in Optimization and Deep Learning*, CSAIL Seminar, MIT, Boston; 31 Oct. 2017, *Adaptation in Artificial Systems: lessons from Evolution Strategies applied to Deep Learning*, XIII Brazilian Congress on Computational Intelligence, Rio de Janeiro.
- Michèle Sebag, July 2017, *AI without hot air / Le vent de l'IA*, Académie des Technologies; July 2017, *IA et Intelligence Service*, DGA Ecole Militaire; Sept. 2017, *Causal Generative Neural Networks*, Lorentz center, Leiden; Sept. 2017, *Stochastic Gradient Descent: Going as fast as possible but not faster*; Sept. 2017, *AutoML@ECMLPKDD*, Skopje; Sept. 2017, *Algorithm Recommender System*, keynote speech JST CREST Program on Big Data Applications, Tokyo.
- Paola Tubaro, 11 May 2017, *Mapping the collaborative economy: social networks, status and norms*, RITM Seminar, Université Paris Sud, Sceaux.

10.1.4. Leadership within the Scientific Community

- Isabelle Guyon, President and co-founder of ChaLearn, a non-for-profit organization dedicated to the organization of challenge.
- Marc Schoenauer, Chair of ACM-SIGEVO (Special Interest Group on Evolutionary Computation), re-elected July 2017 (2-years term).

- Marc Schoenauer, founding President of SPECIES (Society for the Promotion of Evolutionary Computation In Europe and Surroundings), that organizes the yearly series of conferences *EvoStar*.
- Michèle Sebag, elected Chair of Steering Committee, ECML-PKDD; head of the Research Programme, Institut de Convergence DataIA
- Paola Tubaro, convenor of the Social Network Analysis Group of British Sociological Association; co-founder of European Network on Digital Labor

10.1.5. Scientific Expertise

- Cécile Germain, evaluator for the H2020-ICT-2017-1 Big Data PPP call.
- Marc Schoenauer, mission Villani pour l'Intelligence Artificielle

10.1.6. Research Administration

- Cécile Germain, University officer for scientific computing; deputy head of the computer science departement, in charge of research; member of the scientific council of faculty of Science (UPsud) and of its board; member of the Board of the Lidex *Center for Data Science*; member of the scientific council of faculty of Medicine (UPsud).
- Isabelle Guyon, representative of UPSud in the DataIA *Institut de Convergence* Program Committee, University of Paris-Saclay.
- Marc Schoenauer, co-chair (with Sylvain Arlot) of the *Maths-STIC* program of the Labex of Mathematics Hadamard (LMH).
- Michele Sebag, deputy director of LRI, CNRS UMR 8623; elected member of the Research Council of Univ. Paris-Saclay; member of the STIC department council of Univ. Paris-Saclay; member of the Scientific Council of Labex AMIES, Applications des Mathématiques ds l'Industrie, l'Entreprise et la Société; member of the Scientific Council of IRT System'X; member of the CSFRS (Conseil supérieur de la formation et de la recherche stratégique).
- Paola Tubaro, representative of CNRS in the DataIA *Institut de Convergence* Program Committee, University of Paris-Saclay.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence : Philippe Caillou, Computer Science for students in Accounting and Management, 192h, L1, IUT Sceaux, Univ. Paris Sud.

Licence : Aurélien Decelle, Computer Architecture, 28h, L2, Univ. Paris-Sud.

Licence : Aurélien Decelle, Machine Learning and Artificial Life, 55h, L2, Univ. Paris-Sud.

Licence : Aurélien Decelle, Object-oriented programming , 26h, L2, Univ. Paris-Sud.

Licence : Aurélien Decelle, Computer Architecture, 26h, L3, Univ. Paris-Sud.

Licence and Polytech : Cécile Germain, Computer Architecture

Licence : Isabelle Guyon, Project: Creation of mini-challenges, M2, Univ. Paris-Sud.

Master : Guillaume Charpiat and Corentin Tallec, Advanced Machine Learning, 34h, M2 Recherche, Centrale-Supélec.

Master : Aurélien Decelle, Machine Learning, 26h, M1, Univ. Paris-Sud.

Master : Aurélien Decelle, Probability and statistics, 26h, M1, Univ. Paris-Sud.

Master : Cécile Germain, Parallel Programming

Master : Isabelle Guyon, Project: Resolution of mini-challenges (created by M2 students), L2, Univ. Paris-Sud.

Master : Yann Ollivier, Deep learning, 4h, M2 Recherche, Telecom/Polytech.

Master : Michèle Sebag, Machine Learning, 12h; Deep Learning, 6h; Reinforcement Learning, 6h; M2 Recherche, U. Paris-sud.

Master : Paola Tubaro, Sociology of social networks, 24h, M2, EHESS/ENS.

Master : Flora Jay, Population Genetics, 10h, M2, Univ. Paris-Sud.

Doctorate: Paola Tubaro, Research Methods, 12h, University of Insubria, Italy.

10.2.2. Supervision

PhD: Vincent Berthier, *Studies on stochastic optimisation and applications to the real world*, Univ. Paris-Saclay, 29/9/2017, Olivier Teytaud.

PhD: Nacim BELKHIR, *On-line parameter tuning*, Univ. Paris-Saclay, 30/11/2017, Marc Schoenauer and Johann Dréo (Thalès), CIFRE Thalès.

PhD: Pierre-Yves MASSÉ, *Gradient Methods for Statistical Learning*, Univ. Paris-Saclay, 15/12/2017, Yann Ollivier

PhD: Emmanuel MAGGIORI, *Large-Scale Remote Sensing Image Classification*, 22/06/2017, Univ. Nice-Sophia-Antipolis, Yuliya Tarabalka, Pierre Alliez and Guillaume Charpiat

PhD: Yasaman SARABI, *Network Analysis of Private Water Companies, Challenges Collaboration and Competition*, 15/12/2017, Paola Tubaro (at the University of Greenwich, London, UK).

PhD in progress: Mehdi CHERTI *Learning to discover: supervised discrimination and unsupervised representation learning with applications in particle physics.* 01/10/2014, Balazs Kegl.

PhD in progress: Benjamin DONNOT, *Optimisation et méthodes d'apprentissage pour une conduite robuste et efficace du réseau électrique par anticipation sur base de parades topologiques.*, 1/09/2015, Isabelle Guyon and Marc Schoenauer

PhD in progress: Guillaume DOQUET, *ML Algorithm Selection and Domain Adaptation*, 1/09/2015, Michele Sebag

PhD in progress: Victor ESTRADE *Robust domain-adversarial learning, with applications to High Energy Physics,* 01/10/2016, Cécile Germain and Isabelle Guyon.

PhD in progress: François GONARD, *Automatic optimization algorithm selection and configuration*, 1/10/2014, Marc Schoenauer and Michèle Sebag, thèse IRT SystemX.

PhD in progress : Hoang M. LUONG, *Squaring the Circle in Modelling Corporate Governance, Market Structure and Innovation: A Tobin's Q Approach to R&D Investment when Network Effects Are Present*, 01/09/2014, (with M. Ugur and S. Gorgoni, at the University of Greenwich, London, UK).

PhD in progress: Anna PIAZZA, *Inter-Organisational Relationships and Organisational Performance: Network Analysis Applications to a Health Care System*, 01/09/2014, Paola Tubaro (with F. Pallotti and A. Lomi, at the University of Greenwich, London, UK).

PhD in progress: Adrian POL *Machine Learning Anomaly Detection, with application to CMS Data Quality Monitoring,* 01/10/2016, Cécile Germain.

PhD in progress: Thomas SCHMITT, *A Collaborative Filtering Approach to Matching Job Openings and Job Seekers*, 1/11/2014, Philippe Caillou and Michèle Sebag and Jean-Pierre Nadal (EHESS)

PhD in progress: Lisheng SUN, *Apprentissage Automatique: Vers une analyse de données automatisé*, 1/10/2016, Isabelle Guyon and Michèle Sebag

PhD in progress: Corentin TALLEC, *Reinforcement Learning and Recurrent Neural Networks: Dynamical approaches*, 1/10/2016, Yann Ollivier

PhD in progress: Pierre WOLINSKI, *Learning the Architecture of Neural Networks*, 1/9/2016, Guillaume Charpiat and Yann Ollivier

PhD in progress: Victor BERGER, *Variational Anytime Simulator*, 1/10/2017, Michèle Sebag and Marc Schoenauer

PhD in progress: Giancarlo FISSORE, *Statistical physics analysis of generative models*, 1/10/2017, Aurélien Decelle and Cyril Furtlehner

PhD in progress: Diviyan KALAINATHAN, *Causal models and quality of life at work*, 1/10/2017, Michèle Sebag and Isabelle Guyon

PhD in progress: Zhengying LIU, *Automation du design des reseaux de neurones profonds*, 1/10/2017, Isabelle Guyon

PhD in progress: Herilalaina RAKOTOARISON, *Automatic Algorithm Configuration for Power Grid Optimization*, 1/10/2017, Marc Schoenauer and Michèle Sebag

PhD in progress: Théophile SANCHEZ, *Reconstructing the past: deep learning for population genetics*, 1/10/2017, Guillaume Charpiat and Flora Jay

PhD in progress: Aris TRITAS, *Modélisation causale des relations entre alimentation et santé*, 1/10/2017, Michèle Sebag and Philippe Caillou

10.2.3. Juries

Guillaume Charpiat, jury of the 2017 Gilles Kahn PhD prize (SIF); jury of a MdC hiring committee at Univ. Paris-Sud.

Cecile Germain, jury of the Telecom PhD prize; half-way jury of Jacob Montiel (Telecom)

Isabelle Guyon, PhD jury Mathieu Bouyrie (Univ. Paris-Saclay, 11/1/2017) HDR jury Alexandre Gramfort (Univ. Paris-Saclay, 6/11/2017)

Flora Jay, half-way juries of Bérénice Alard (MNHN), Arnaud Becheler (EGCE), Cyriel Paris (INRA Toulouse)

Marc Schoenauer, PhD jury of Mathieu Carriere (Univ. Paris-Saclay, 21/11/2017), Elvis Dohmatob (Univ. Paris-Saclay, 29/9/2017) ; PhD committee and half-way jury of Arthur Mensch (Univ. Paris-Saclay, 3/7/2017), Julio Navarro Lara (Univ. Strasbourg, 25/7/2017), Jean Marçais (Univ. Rennes, 10/2017).

Michèle Sebag, Research Quality Assessment Panel for the Department of Computer Science, U. Copenhagen; Hiring Jury, Professor U. Nice Cote d'Azur; Hiring Jury, Professor U. Dortmund, Germany; Hiring Jury, MdC UPSud; Reviewer PhD Audrey Durand, U. Laval Québec; Reviewer PhD Antonio Vergari, U. Bari, Italie; Jury Member: Gisselbrecht, UPMC; Renyu Xury, UPSud.

10.3. Popularization

Philippe Caillou, talk on quality of life at work, AFRAME association for ethic management, Paris, 28/11/2017

Isabelle Guyon: press release at NIPS conference (5/12/2015); interview l'Usine Nouvelle (12/12/2017).

Michèle Sebag: interview France 2; interview l'Usine Nouvelle; 2 articles The Conversation France.

Paola Tubaro: interview Pourquoi Docteur (online health magazine, 01/01/2017); article in Journal du CNRS (02/02/2017); op-ed Libération (02/02/2017); interview Le Devoir (daily newspaper, Canada, 29/04/2017); interview Les Echos (19/09/2017); interview Le Monde (09/11/2017); 1 article OuiShare magazine (13/07/2017); 1 article The Conversation France (French version 02/11/2017, English version 11/12/2017); round table on Open data at "Science and You" conference, Montréal (05/05/2017); round table "Cuisine and Performance" at Centre Pompidou, Paris (18/11/2017).

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] N. BELKHIR.*Per Instance Algorithm Configuration for Continuous Black Box Optimization*, Paris Saclay, November 2017, https://hal.inria.fr/tel-01669527.

[2] V. BERTHIER.*Studies on stochastic optimisation and applications to the real world*, Université Paris 11, September 2017, https://hal.inria.fr/tel-01668371.

[3] P.-Y. MASSÉ.*Around the Use of Gradients in Machine Learning*, Université Paris-Saclay, December 2017, https://hal.archives-ouvertes.fr/tel-01665478.

Articles in International Peer-Reviewed Journal

[4] K. CAYE, F. JAY, O. MICHEL, O. FRANÇOIS.*Fast Inference of Individual Admixture Coefficients Using Geographic Data*, in "Annals Of Applied Statistics", 2018, https://hal.archives-ouvertes.fr/hal-01676712.

[5] A. DECELLE, G. FISSORE, C. FURTLEHNER.*Spectral Dynamics of Learning Restricted Boltzmann Machines*, in "EPL - Europhysics Letters", November 2017, https://hal.inria.fr/hal-01652314.

[6] C. FERNANDO CRISPIM-JUNIOR, A. GÓMEZ URÍA, C. STRUMIA, M. KOPERSKI, A. KONIG, F. NEGIN, S. COSAR, A.-T. NGHIEM, G. CHARPIAT, F. BREMOND, D. P. CHAU.*Online recognition of daily activities by color-depth sensing and knowledge models*, in "Sensors", June 2017, vol. 17, n^{o} 7, p. 1-15 [*DOI :* 10.3390/s17071528], https://hal.inria.fr/hal-01658438.

[7] I. GUYON, H. J. ESCALANTE, V. ATHITSOS, P. JANGYODSUK, J. WAN.*Principal motion components for one-shot gesture recognition*, in "Pattern Analysis and Applications", February 2017, vol. 20, n^{o} 1, p. 167 - 182 [*DOI :* 10.1007/s10044-015-0481-3], https://hal.inria.fr/hal-01677941.

[8] Y. GÜÇLÜTÜRK, U. GÜÇLÜ, X. BARÓ, H. J. ESCALANTE, I. GUYON, S. ESCALERA, M. A. J. VAN GERVEN, R. VAN LIER.*Multimodal First Impression Analysis with Deep Residual Networks*, in "IEEE Transactions on Affective Computing", September 2017, vol. PP, n^{o} 99, p. 1-14 [*DOI :* 10.1109/TAFFC.2017.2751469], https://hal.inria.fr/hal-01668375.

[9] E. MAGGIORI, Y. TARABALKA, G. CHARPIAT, P. ALLIEZ.*High-Resolution Semantic Labeling with Convolutional Neural Networks*, in "IEEE Transactions on Geoscience and Remote Sensing", December 2017, https://arxiv.org/abs/1611.01962 , https://hal.inria.fr/hal-01393279.

[10] M. MISIR, M. SEBAG.*Alors: An algorithm recommender system*, in "Artificial Intelligence", March 2017, vol. 244, p. 291-314, Published on-line Dec. 2016, https://hal.inria.fr/hal-01419874.

[11] Y. OLLIVIER, L. ARNOLD, A. AUGER, N. HANSEN.*Information-Geometric Optimization Algorithms: A Unifying Picture via Invariance Principles*, in "Journal of Machine Learning Research", 2017, vol. 18, n^{o} 18, p. 1-65, https://hal.inria.fr/hal-01515898.

[12] F. PALLOTTI, P. TUBARO, A. A. CASILLI, T. W. VALENTE.*"You see yourself like in a mirror": The effects of internet-mediated personal networks on body image and eating disorders*, in "Health Communication", 2017, Published online on 6 July 2017 [*DOI :* 10.1080/10410236.2017.1339371], https://hal.archives-ouvertes.fr/hal-01520138.

[13] P. TUBARO.*Les tensions entre sociologie et politique à l'aune d'une tentative de législation des sites web sur les troubles alimentaires* , in "SociologieS", November 2017, p. 1-13, https://hal.archives-ouvertes.fr/hal-01648305.

Invited Conferences

[14] F. JAY, S. S. BOITARD, F. AUSTERLITZ.*Reconstructing past history from whole-genomes: an ABC approach handling recombining data*, in "European Mathematical Genetics Meeting", Tartu, Estonia, April 2017, https://hal.archives-ouvertes.fr/hal-01679379.

International Conferences with Proceedings

[15] E. AGUSTSSON, R. TIMOFTE, S. ESCALERA, X. BARÓ, I. GUYON, R. ROTHE.*Apparent and real age estimation in still images with deep residual regressors on APPA-REAL database*, in "FG 2017 - 12th IEEE International Conference on Automatic Face and Gesture Recognition", Washington DC, United States, May 2017, p. 1-12, https://hal.inria.fr/hal-01677892.

[16] S. AMROUCHE, N. BRAUN, P. CALAFIURA, S. FARRELL, J. GEMMLER, C. GERMAIN, V. GLIGOROV, T. GOLLING, H. GRAY, I. GUYON, M. HUSHCHYN, V. INNOCENTE, B. KÉGL, S. NEUHAUS, D. ROUSSEAU, A. SALZBURGER, A. USTYUZHANIN, J.-R. VLIMANT, C. WESSEL, Y. YILMAZ.*Track reconstruction at LHC as a collaborative data challenge use case with RAMP*, in "Connecting The Dots / Intelligent Tracker", Orsay, France, March 2017, vol. 150, 00015 [*DOI :* 10.1051/EPJCONF/201715000015], https://hal.archives-ouvertes.fr/hal-01584689.

[17] M. ASADI-AGHBOLAGHI, A. CLAPES, M. BELLANTONIO, H. J. ESCALANTE, V. PONCE-LÓPEZ, X. BARÓ, I. GUYON, S. KASAEI, S. ESCALERA.*A Survey on Deep Learning Based Approaches for Action and Gesture Recognition in Image Sequences*, in "FG 2017 - 12th IEEE Conference on Automatic Face and Gesture Recognition", Washington, DC, United States, IEEE, May 2017, p. 476-483 [*DOI :* 10.1109/FG.2017.150], https://hal.inria.fr/hal-01668383.

[18] N. BELKHIR, J. DREO, P. SAVÉANT, M. SCHOENAUER.*Per instance algorithm configuration of CMA-ES with limited budget*, in "GECCO 2017 - Proceedings of the Genetic and Evolutionary Computation Conference", Berlin, Germany, July 2017, p. 681-688, https://hal.inria.fr/hal-01613753.

[19] I. BRIGUI-CHTIOUI, P. CAILLOU, E. NEGRE.*Intelligent Digital Learning: Agent-Based Recommender System*, in "ICMLC 2017 - 9th International Conference on Machine Learning and Computing", Singapore, Singapore, February 2017 [*DOI :* 10.1145/3055635.3056592], https://hal.inria.fr/hal-01680527.

[20] B. DONNOT, I. GUYON, M. SCHOENAUER, P. PANCIATICI, A. MAROT.*Introducing machine learning for power system operation support*, in "IREP Symposium", Espinho, Portugal, August 2017, https://arxiv.org/abs/1709.09527 , https://hal.inria.fr/hal-01581719.

[21] H. J. ESCALANTE, I. GUYON, S. ESCALERA, J. JACQUES, M. MADADI, X. BARÓ, S. AYACHE, E. VIEGAS, Y. GÜÇLÜTÜRK, U. GÜÇLÜ, M. A. J. VAN GERVEN, R. VAN LIER.*Design of an Explainable Machine Learning Challenge for Video Interviews*, in "IJCNN 2017 - 30th International Joint Conference on Neural Networks", Anchorage, AK, United States, Neural Networks (IJCNN), 2017 International Joint Conference on, IEEE, May 2017, p. 1-8 [*DOI :* 10.1109/IJCNN.2017.7966320], https://hal.inria.fr/hal-01668386.

[22] E. GALVÁN-LÓPEZ, L. VA´ZQUEZ-MENDOZA, M. SCHOENAUER, L. TRUJILLO.*On the Use of Dynamic GP Fitness Cases in Static and Dynamic Optimisation Problems*, in "EA 2017- International Conference on Artificial Evolution", Paris, France, Evelyne Lutton, October 2017, p. 1-14, https://hal.inria.fr/hal-01648365.

[23] F. GONARD, M. SCHOENAUER, M. SEBAG.*ASAP.V2 and ASAP.V3: Sequential optimization of an Algorithm Selector and a Scheduler*, in "Open Algorithm Selection Challenge 2017", Brussels, Belgium, Proceedings of Machine Learning Research, September 2017, vol. 79, p. 8-11, https://hal.inria.fr/hal-01659700.

[24] Y. GÜÇLÜTÜRK, U. GÜÇLÜ, M. PEREZ, H. J. ESCALANTE BALDERAS, X. BARÓ, I. GUYON, C. ANDUJAR, J. J. JUNIOR, M. MADADI, S. ESCALERA, M. A. J. VAN GERVEN, R. VAN LIER.*Visualizing Apparent Personality Analysis with Deep Residual Networks*, in "International Conference on Computer Vision - ICCV 2017", Venice, Italy, October 2017, https://hal.inria.fr/hal-01677962.

[25] E. MAGGIORI, Y. TARABALKA, G. CHARPIAT, P. ALLIEZ.*Can Semantic Labeling Methods Generalize to Any City? The Inria Aerial Image Labeling Benchmark*, in "IEEE International Symposium on Geoscience and Remote Sensing (IGARSS)", Fort Worth, United States, July 2017, https://hal.inria.fr/hal-01468452.

[26] E. MAGGIORI, Y. TARABALKA, G. CHARPIAT, P. ALLIEZ.*Polygonization of Remote Sensing Classification Maps by Mesh Approximation*, in "ICIP 2017 - IEEE International Conference on Image Processing", Beijing, China, September 2017, 5, https://hal.inria.fr/hal-01530460.

[27] L. MARTÍ, A. FANSI-TCHANGO, L. NAVARRO, M. SCHOENAUER.*Progressively Adding Objectives: A Case Study in Anomaly Detection*, in "Genetic and Evolutionary Computation Conference (GECCO 2017)", Berlin, Germany, July 2017 [*DOI :* 10.1145/3071178.3071333], https://hal.inria.fr/hal-01525611.

[28] *Best Paper*
Y. OLLIVIER, G. MARCEAU-CARON.*Natural Langevin Dynamics for Neural Networks*, in "GSI 2017 - 3rd conference on Geometric Science of Information", Paris, France, Springer Verlag, November 2017, vol. 10589, p. 451-459, https://arxiv.org/abs/1712.01076 - Best Paper Award [*DOI :* 10.1007/978-3-319-68445-1_53], https://hal.archives-ouvertes.fr/hal-01655949.

[29] T. SCHMITT, F. GONARD, P. CAILLOU, M. SEBAG.*Language Modelling for Collaborative Filtering: Application to Job Applicant Matching*, in "ICTAI 2017 - 29th IEEE International Conference on Tools with Artificial Intelligence", Boston, United States, November 2017, p. 1-8, https://hal.inria.fr/hal-01659543.

[30] J. WAN, S. ESCALERA, G. ANBARJAFARI, H. J. ESCALANTE, X. BARÓ, I. GUYON, M. MADADI, J. ALLIK, C. LIN, Y. XIE.*Results and Analysis of ChaLearn LAP Multi-modal Isolated and Continuous Gesture Recognition, and Real versus Fake Expressed Emotions Challenges*, in "International Conference on Computer Vision - ICCV 2017", Venice, Italy, October 2017, https://hal.inria.fr/hal-01677974.

[31] N. YOSHIKAWA, N. BELKHIR, S. SUZUKI.*Recurrent Neural Network-based Fault Detector for Aileron Failures of Aircraft*, in "ASCC 2017 - The 2017 Asian Control Conference", Gold Coast, Australia, December 2017, https://hal.inria.fr/hal-01669540.

Conferences without Proceedings

[32] V. ESTRADE, C. GERMAIN, I. GUYON, D. ROUSSEAU.*Adversarial learning to eliminate systematic errors: a case study in High Energy Physics*, in "NIPS 2017 - workshop Deep Learning for Physical Sciences", Long Beach, United States, December 2017, p. 1-5, https://hal.inria.fr/hal-01665925.

[33] T. SANCHEZ, G. CHARPIAT, F. JAY.*SPI-DNA: End-to-end Deep Learning Approach for Demographic History Inference*, in "Paris-Saclay Junior Conference on Data Science and Engineering", Orsay, France, September 2017, https://hal.archives-ouvertes.fr/hal-01679385.

Scientific Books (or Scientific Book chapters)

[34] M. ASADI-AGHBOLAGHI, A. CLAPES, M. BELLANTONIO, H. J. ESCALANTE, V. PONCE-LÓPEZ, X. BARÓ, I. GUYON, S. KASAEI, S. ESCALERA.*Deep learning for action and gesture recognition in image sequences: a survey* , in "Gesture Recognition", Springer Verlag, July 2017, p. 539-578, A reduced version of this paper appeared appeared in the Proceedings of 12th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2017), 2017, https://hal.inria.fr/hal-01678006.

[35] F. GONARD, M. SCHOENAUER, M. SEBAG.*Algorithm Selector and Prescheduler in the ICON challenge*, in "Bioinspired heuristic optimization", E.-G. TALBI, A. NAKIB (editors), Computational Intelligence, Springer Verlag, 2017, https://hal.inria.fr/hal-01663174.

[36] D. KALAINATHAN, O. GOUDET, P. CAILLOU, P. TUBARO, T. WEIL, E. BOURDU.*Portraits de travailleurs : Comprendre la qualité de vie au travail*, Presses des Mines - La Fabrique de l'Industrie, 2017, https://hal.inria.fr/hal-01665423.

Research Reports

[37] A. DECELLE, G. FISSORE, C. FURTLEHNER.*Thermodynamics of Restricted Boltzmann Machines and Related Learning Dynamics*, Inria Saclay Ile de France ; LRI, Université Paris-Sud, January 2018, n^{o} RR-9139, p. 1-36, https://hal.inria.fr/hal-01675310.

Scientific Popularization

[38] A. A. CASILLI, P. TUBARO.*Réprimer les sites «pro-anorexie» : une fausse bonne idée* , in "Libération", February 2017, p. 1-2, https://hal.archives-ouvertes.fr/hal-01577074.

[39] P. TUBARO.*Social Networks: face-to-face and online ties at OuiShare Fest*, July 2017, Texte court publié dans OuiShare Magazine, https://hal.archives-ouvertes.fr/hal-01666817.

Other Publications

[40] O. BOUSQUET, S. GELLY, K. KURACH, M. SCHOENAUER, M. SEBAG, O. TEYTAUD, D. VINCENT.*Toward Optimal Run Racing: Application to Deep Learning Calibration*, June 2017, https://arxiv.org/abs/1706.03199 - working paper or preprint, https://hal.inria.fr/hal-01634381.

[41] A. A. CASILLI, P. TUBARO.*Rethinking ethics in social-network research*, December 2017, Texte de vulgarisation publié dans The Conversation, une version française existe également, https://hal.archives-ouvertes.fr/hal-01666784.

[42] B. DONNOT, I. GUYON, M. SCHOENAUER, A. MAROT, P. PANCIATICI.*Fast Power system security analysis with Guided Dropout, supplemental material*, November 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01649938.

[43] V. ESTRADE, C. GERMAIN, I. GUYON, D. ROUSSEAU.*Robust deep learning: A case study*, September 2017, p. 1-5, JDSE 2017 - 2nd Junior Conference on Data Science and Engineering, https://hal.inria.fr/hal-01665938.

[44] C. GERMAIN.*Summary of the Weizmann workshop: Hammers & Nails - Machine Learning & HEP*, July 2017, p. 1-48, 2017 - Hammers & Nails - Machine Learning & HEP, https://hal.inria.fr/hal-01665940.

[45] O. GOUDET, D. KALAINATHAN, P. CAILLOU, D. LOPEZ-PAZ, I. GUYON, M. SEBAG, A. TRITAS, P. TUBARO.*Learning Functional Causal Models with Generative Neural Networks*, November 2017, https://arxiv.org/abs/1709.05321 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01649153.

[46] I. GUYON, S. ESCALERA, X. BARÓ, H. J. ESCALANTE.*ChaLearn Looking at People: A Review of Events and Resources*, January 2018, https://arxiv.org/abs/1701.02664 - Paper to appear in proceedings of IJCNN 2017 - IEEE - Associated website: http://chalearnlap.cvc.uab.es, https://hal.inria.fr/hal-01677944.

[47] A. MAROT, S. TAZI, B. DONNOT, P. PANCIATICI.*Large-scale power grid hierarchical segmentation based on power-flow affinities*, November 2017, https://arxiv.org/abs/1711.09715 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01633508.

[48] Y. OLLIVIER.*Online Natural Gradient as a Kalman Filter*, December 2017, https://arxiv.org/abs/1703.00209 - working paper or preprint, https://hal.inria.fr/hal-01660622.

[49] Y. OLLIVIER, C. TALLEC.*Unbiased Online Recurrent Optimization*, December 2017, https://arxiv.org/abs/1702.05043 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01666483.

[50] Y. OLLIVIER, C. TALLEC.*Unbiasing Truncated Backpropagation Through Time*, December 2017, https://arxiv.org/abs/1705.08209 - working paper or preprint, https://hal.inria.fr/hal-01660627.

[51] D. ROUSSEAU, S. AMROUCHE, P. CALAFIURA, S. FARRELL, C. GERMAIN, V. GLIGOROV, T. GOLLING, H. GRAY, I. GUYON, M. HUSHCHYN, V. INNOCENTE, M. KIEHN, A. SALZBURGER, A. USTYUZHANIN, J.-R. V. VLIMANT, Y. YILMAZ.*WCCI 2018 TrackML Particle Tracking Challenge*, July 2018, The document describes the challenge data, task and organization, https://hal.inria.fr/hal-01680537.

[52] A. SCHOENAUER-SEBAG, M. SCHOENAUER, M. SEBAG.*Stochastic Gradient Descent: Going As Fast As Possible But Not Faster*, September 2017, https://arxiv.org/abs/1709.01427 - working paper or preprint, https://hal.inria.fr/hal-01634375.

[53] N. B. SHAH, B. TABIBIAN, K. MUANDET, I. GUYON, U. VON LUXBURG.*Design and Analysis of the NIPS 2016 Review Process*, December 2017, https://arxiv.org/abs/1708.09794 - working paper or preprint, https://hal.inria.fr/hal-01668377.

References in notes

[54] S. ESCALERA, I. GUYON, V. ATHITSOS (editors). *Gesture Recognition*, Springer, 2017, https://doi.org/10.1007/978-3-319-57021-1.

[55] C. ADAM-BOURDARIOS, G. COWAN, C. GERMAIN, I. GUYON, B. KÉGL, D. ROUSSEAU.*How Machine Learning won the Higgs Boson Challenge*, in "Proc. European Symposium on ANN, CI and ML", 2016.

[56] R. AKROUR, M. SCHOENAUER, M. SEBAG.*Programming by Feedback*, in "International Conference on Machine Learning", JMLR Proceedings, JMLR.org, 2014, n° 32, p. 1503-1511.

[57] M. AMIL, N. BREDÈCHE, C. GAGNÉ, S. GELLY, M. SCHOENAUER, O. TEYTAUD.*A Statistical Learning Perspective of Genetic Programming*, in "EuroGP", Proceedings of EuroGP 09, Springer Verlag, 2009.

[58] M. ANDRYCHOWICZ, M. DENIL, S. GÓMEZ, M. W. HOFFMAN, D. PFAU, T. SCHAUL, N. DE FREITAS.*Learning to learn by gradient descent by gradient descent*, in "Advances in Neural Information Processing Systems 29", D. D. LEE, M. SUGIYAMA, U. V. LUXBURG, I. GUYON, R. GARNETT (editors), Curran Associates, Inc., 2016, p. 3981–3989.

[59] J. BERGSTRA, R. BARDENET, Y. BENGIO, B. KÉGL.*Algorithms for Hyper-Parameter Optimization*, in "25th Annual Conference on Neural Information Processing Systems (NIPS 2011)", J. SHAWE-TAYLOR, R. ZEMEL, P. BARTLETT, F. PEREIRA, K. WEINBERGER (editors), Advances in Neural Information Processing Systems, Neural Information Processing Systems Foundation, 2011, vol. 24.

[60] S. BOITARD, W. RODRÍGUEZ, F. JAY, S. MONA, F. AUSTERLITZ.*Inferring population size history from large samples of genome-wide molecular data-an approximate Bayesian computation approach*, in "PLoS genetics", 2016, vol. 12, n^{o} 3, e1005877.

[61] P. BRAZDIL, C. GIRAUD CARRIER, C. SOARES, R. VILALTA.*Metalearning: Applications to Data Mining*, Springer Verlag, 2009.

[62] E. BROCHU, V. M. CORA, N. DE FREITAS.*A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning*, in "arXiv preprint arXiv:1012.2599", 2010.

[63] S. BRUNTON, J. PROCTOR, N. KUTZ.*Sparse Identification of Nonlinear Dynamics (SINDy)*, in "Bulletin of the American Physical Society", 2016, vol. 61.

[64] T. DOMHAN, J. T. SPRINGENBERG, F. HUTTER.*Speeding Up Automatic Hyperparameter Optimization of Deep Neural Networks by Extrapolation of Learning Curves*, in "Proceedings of the 24th International Conference on Artificial Intelligence", IJCAI'15, AAAI Press, 2015, p. 3460–3468, http://dl.acm.org/citation.cfm?id=2832581.2832731.

[65] B. DONNOT, I. GUYON, M. SCHOENAUER, A. MAROT, P. PANCIATICI.*Fast Power system security analysis with Guided Dropout*, 2018, submitted.

[66] C. FURTLEHNER, A. DECELLE.*Cycle-based Cluster Variational Method for Direct and Inverse Inference*, in "Journal of Statistical Physics", 2016, vol. 164, n^{o} 3, p. 531–574.

[67] Y. GANIN, E. USTINOVA, H. AJAKAN, P. GERMAIN, H. LAROCHELLE, F. LAVIOLETTE, M. MARCHAND, V. LEMPITSKY.*Domain-Adversarial Training of Neural Networks*, in "Journal of Machine Learning Research", 2016, vol. 17, n^{o} 59, p. 1-35.

[68] I. GOODFELLOW, J. POUGET-ABADIE, M. MIRZA, B. XU, D. WARDE-FARLEY, S. OZAIR, A. COURVILLE, Y. BENGIO.*Generative Adversarial Nets*, in "NIPS 27", Z. GHAHRAMANI, M. WELLING, C. CORTES, N. D. LAWRENCE, K. Q. WEINBERGER (editors), Curran Associates, Inc., 2014, p. 2672–2680.

[69] I. GUYON, C. F. ALIFERIS, G. F. COOPER, A. ELISSEEFF, J.-P. PELLET, P. SPIRTES, A. R. STATNIKOV.*Design and Analysis of the Causation and Prediction Challenge*, in "WCCI Causation and Prediction Challenge", JMLR W-CP, 2008, p. 1–33.

[70] I. GUYON, K. BENNETT, G. CAWLEY, H. J. ESCALANTE, S. ESCALERA, T. K. HO, N. MACIA, B. RAY, M. SAEED, A. STATNIKOV.*Design of the 2015 ChaLearn AutoML challenge*, in "Proc. IJCNN", IEEE, 2015, p. 1–8.

[71] H. H. HOOS.*Programming by Optimization*, in "Commun. ACM", 2012, vol. 55, n° 2, p. 70–80.

[72] Y. LECUN, Y. BENGIO, G. HINTON.*Deep learning*, in "Nature", 2015, vol. 521, n° 7553, p. 436–444.

[73] E. MAGGIORI.*Learning approaches for large-scale remote sensing image classification*, Université Côte d'Azur, June 2017, https://hal.inria.fr/tel-01589661.

[74] L. MARTI, A. FANSI-TCHANGO, L. NAVARRO, M. SCHOENAUER.*Anomaly Detection with the Voronoi Diagram Evolutionary Algorithm*, in "Parallel Problem Solving from Nature – PPSN XIV", Edinburgh, United Kingdom, J. HANDL, E. HART, P. LEWIS, M. LÓPEZ-IBÁÑEZ, G. OCHOA, B. PAECHTER (editors), LNCS, Springer Verlag, September 2016, vol. 9921, p. 697-706 [*DOI :* 10.1007/978-3-319-45823-6_65], https://hal.inria.fr/hal-01387621.

[75] J. M. MOOIJ, J. PETERS, D. JANZING, J. ZSCHEISCHLER, B. SCHÖLKOPF.*Distinguishing Cause from Effect Using Observational Data: Methods and Benchmarks*, in "Journal of Machine Learning Research", 2016, vol. 17, n° 32, p. 1-102.

[76] Y. OLLIVIER.*Riemannian metrics for neural networks I: Feedforward networks*, in "Information and Inference", 2015, vol. 4, n° 2, p. 108–153.

[77] Y. OLLIVIER.*Riemannian metrics for neural networks II: Recurrent networks and learning symbolic data sequences*, in "Information and Inference", 2015, vol. 4, n° 2, p. 154–193.

[78] J. PEARL.*Causality: Models, Reasoning, and Inference (2nd edition)*, Cambridge University Press, 2009.

[79] D. ROUSSEAU, P. CALAFIURA, C. GERMAIN, V. INNOCENTE, R. CENCI, M. KAGAN, I. GUYON, D. CLARK, S. FARREL, R. CARNEY, A. SALZBURGER, D. COSTANZO, M. ELSING, T. GOLLING, T. TONG, J.-R. V. VLIMANT.*TrackML: a LHC Tracking Machine Learning Challenge*, in "International Conference on Computing in High Energy and Nuclear Physics", San Francisco, United States, October 2016, https://hal.inria.fr/hal-01422939.

[80] I. SHPITSER, K. MOHAN, J. PEARL.*Missing Data as a Causal and Probabilistic Problem*, in "Proceedings of the Thirty-First Conference on Uncertainty in Artificial Intelligence, UAI 2015, July 12-16, 2015, Amsterdam, The Netherlands", M. MEILA, T. HESKES (editors), AUAI Press, 2015, p. 802–811, http://auai.org/uai2015/proceedings/papers/204.pdf.

[81] D. H. STERN, H. SAMULOWITZ, R. HERBRICH, T. GRAEPEL, L. PULINA, A. TACCHELLA.*Collaborative Expert Portfolio Management*, in "Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2010", M. FOX, D. POOLE (editors), AAAI Press, 2010.

[82] C. THORNTON, F. HUTTER, H. H. HOOS, K. LEYTON-BROWN.*Auto-WEKA: Combined Selection and Hyperparameter Optimization of Classification Algorithms*, in "Proc. of KDD-2013", 2013, p. 847-855.

[83] S. TRIANTAFILLOU, I. TSAMARDINOS.*Constraint-based Causal Discovery from Multiple Interventions over Overlapping Variable Sets*, in "Journal of Machine Learning Research", 2015, vol. 16, p. 2147-2205.

[84] M. YAGOUBI, M. SCHOENAUER.*Asynchronous Master/Slave MOEAs and Heterogeneous Evaluation Costs*, in "Genetic and Evolutionary Computation Conference (GECCO 2012)", United States, T. SOULE, J. H. MOORE (editors), ACM Press, July 2012, p. 1007-1014, https://hal.archives-ouvertes.fr/hal-00689965.

[85] M. YAGOUBI.*Multi-objective parallel evolutionary algorithms : Application to Diesel Combustion*, Université Paris Sud - Paris XI, July 2012, https://tel.archives-ouvertes.fr/tel-00734108.

Project-Team TOCCATA

Certified Programs, Certified Tools, Certified Floating-Point Computations

IN COLLABORATION WITH: Laboratoire de recherche en informatique (LRI)

IN PARTNERSHIP WITH:
Université Paris-Sud (Paris 11)

RESEARCH CENTER
Saclay - Île-de-France

THEME
Proofs and Verification

Table of contents

Project-Team TOCCATA

Creation of the Team: 2012 September 01, updated into Project-Team: 2014 July 01

Keywords:

Computer Science and Digital Science:

A2.1.1. - Semantics of programming languages
A2.1.3. - Functional programming
A2.1.6. - Concurrent programming
A2.1.10. - Domain-specific languages
A2.1.11. - Proof languages
A2.4.2. - Model-checking
A2.4.3. - Proofs
A6.2.1. - Numerical analysis of PDE and ODE
A7.2. - Logic in Computer Science
A7.2.1. - Decision procedures
A7.2.2. - Automated Theorem Proving
A7.2.3. - Interactive Theorem Proving
A7.2.4. - Mechanized Formalization of Mathematics
A8.10. - Computer arithmetic

Other Research Topics and Application Domains:

B5.2.2. - Railway
B5.2.3. - Aviation
B5.2.4. - Aerospace
B6.1. - Software industry
B9.4.1. - Computer science
B9.4.2. - Mathematics

1. Personnel

Research Scientists

Claude Marché [Team leader, Inria, Senior Researcher, HDR]
Sylvie Boldo [Inria, Senior Researcher, HDR]
Jean-Christophe Filliâtre [CNRS, Senior Researcher, HDR]
Guillaume Melquiond [Inria, Researcher]

Faculty Members

Sylvain Conchon [Univ Paris-Sud, Professor, HDR]
Andrei Paskevich [Univ Paris-Sud, Associate Professor]
Thibault Hilaire [Associate Professor, Délégation de l'Univ Pierre et Marie Curie, from Sep 2017]

Technical Staff

Sylvain Dailler [Inria]
Clément Fumex [Inria, until Mar 2017]

PhD Students

Ran Chen [Institute of Software, Chinese Academy of Sciences, Beijing, China, visiting until Feb. 2017]

Martin Clochard [Univ Paris-Sud until Aug, Inria since Sep]
Albin Coquereau [École Nationale Supérieure de Techniques Avancées]
David Declerck [Univ Paris-Sud]
Florian Faissole [Inria]
Diane Gallois-Wong [Univ Paris-Sud, from Oct 2017]
Mário Pereira [Grant Portuguese government]
Raphaël Rieu-Helft [CIFRE TrustInSoft, from Oct 2017]
Mattias Roux [Univ Paris-Sud]

Administrative Assistant

Katia Evrat [Inria]

2. Overall Objectives

2.1. Overall Objectives

The general objective of the Toccata project is to promote formal specification and computer-assisted proof in the development of software that requires high assurance in terms of safety and correctness with respect to the intended behavior of the software.

2.1.1. Context

The importance of software in critical systems increased a lot in the last decade. Critical software appears in various application domains like transportation (e.g., aviation, railway), communication (e.g., smartphones), banking, etc. The number of tasks performed by software is quickly increasing, together with the number of lines of code involved. Given the need of high assurance of safety in the functional behavior of such applications, the need for automated (i.e., computer-assisted) methods and techniques to bring guarantee of safety became a major challenge. In the past and at present, the most widely used approach to check safety of software is to apply heavy test campaigns. These campaigns take a large part of the costs of software development, yet they cannot ensure that all the bugs are caught.

Generally speaking, software verification approaches pursue three goals: (1) verification should be sound, in the sense that no bugs should be missed, (2) verification should not produce false alarms, or as few as possible (3) it should be as automated as possible. Reaching all three goals at the same time is a challenge. A large class of approaches emphasizes goals (2) and (3): testing, run-time verification, symbolic execution, model checking, etc. Static analysis, such as abstract interpretation, emphasizes goals (1) and (3). Deductive verification emphasizes (1) and (2). The Toccata project is mainly interested in exploring the deductive verification approach, although we also consider the others in some cases.

In the past decade, there has been significant progress made in the domain of deductive program verification. They are emphasized by some success stories of application of these techniques on industrial-scale software. For example, the *Atelier B* system was used to develop part of the embedded software of the Paris metro line 14 [50] and other railroad-related systems; a formally proved C compiler was developed using the Coq proof assistant [110]; Microsoft's hypervisor for highly secure virtualization was verified using VCC [89] and the Z3 prover [129]; the L4-verified project developed a formally verified micro-kernel with high security guarantees, using analysis tools on top of the Isabelle/HOL proof assistant [106]. Another sign of recent progress is the emergence of deductive verification competitions (e.g., VerifyThis [2], VScomp [100]).

Finally, recent trends in the industrial practice for development of critical software is to require more and more guarantees of safety, e.g., the upcoming DO-178C standard for developing avionics software adds to the former DO-178B the use of formal models and formal methods. It also emphasizes the need for certification of the analysis tools involved in the process.

2.1.2. Deductive verification

There are two main families of approaches for deductive verification. Methods in the first family build on top of mathematical proof assistants (e.g., Coq, Isabelle) in which both the model and the program are encoded; the proof that the program meets its specification is typically conducted in an interactive way using the underlying proof construction engine. Methods from the second family proceed by the design of standalone tools taking as input a program in a particular programming language (e.g., C, Java) specified with a dedicated annotation language (e.g., ACSL [49], JML [71]) and automatically producing a set of mathematical formulas (the *verification conditions*) which are typically proved using automatic provers (e.g., Z3, Alt-Ergo [52], CVC3 [48], CVC4).

The first family of approaches usually offers a higher level of assurance than the second, but also demands more work to perform the proofs (because of their interactive nature) and makes them less easy to adopt by industry. Moreover, they do not allow to directly analyze a program written in a mainstream programming language like Java or C. The second kind of approaches has benefited in the past years from the tremendous progress made in SAT and SMT solving techniques, allowing more impact on industrial practices, but suffers from a lower level of trust: in all parts of the proof chain (the model of the input programming language, the VC generator, the back-end automatic prover), potential errors may appear, compromising the guarantee offered. Moreover, while these approaches are applied to mainstream languages, they usually support only a subset of their features.

3. Research Program

3.1. Introduction

In the former ProVal project, we have been working on the design of methods and tools for deductive verification of programs. One of our original skills was the ability to conduct proofs by using automatic provers and proof assistants at the same time, depending on the difficulty of the program, and specifically the difficulty of each particular verification condition. We thus believe that we are in a good position to propose a bridge between the two families of approaches of deductive verification presented above. Establishing this bridge is one of the goals of the Toccata project: we want to provide methods and tools for deductive program verification that can offer both a high amount of proof automation and a high guarantee of validity. Toward this objective, a new axis of research was proposed: the development of *certified* analysis tools that are themselves formally proved correct.

The reader should be aware that the word “certified” in this scientific programme means “verified by a formal specification and a formal proof that the program meets this specification”. This differs from the standard meaning of “certified” in an industrial context where it means a conformance to some rigorous process and/or norm. We believe this is the right term to use, as it was used for the *Certified Compiler* project [110], the new conference series *Certified Programs and Proofs*, and more generally the important topics of *proof certificates*.

In industrial applications, numerical calculations are very common (e.g. control software in transportation). Typically they involve floating-point numbers. Some of the members of Toccata have an internationally recognized expertise on deductive program verification involving floating-point computations. Our past work includes a new approach for proving behavioral properties of numerical C programs using Frama-C/Jessie [46], various examples of applications of that approach [68], the use of the Gappa solver for proving numerical algorithms [128], an approach to take architectures and compilers into account when dealing with floating-point programs [69], [121]. We also contributed to the Handbook of Floating-Point Arithmetic [120]. A representative case study is the analysis and the proof of both the method error and the rounding error of a numerical analysis program solving the one-dimension acoustic wave equation [3] [60]. Our experience led us to a conclusion that verification of numerical programs can benefit a lot from combining automatic and interactive theorem proving [62], [68]. Certification of numerical programs is the other main axis of Toccata.

Our scientific programme in structured into four objectives:

1. deductive program verification;
2. automated reasoning;
3. formalization and certification of languages, tools and systems;
4. proof of numerical programs.

We detail these objectives below.

3.2. Deductive Program Verification

Permanent researchers: A. Charguéraud, S. Conchon, J.-C. Filliâtre, C. Marché, G. Melquiond, A. Paskevich

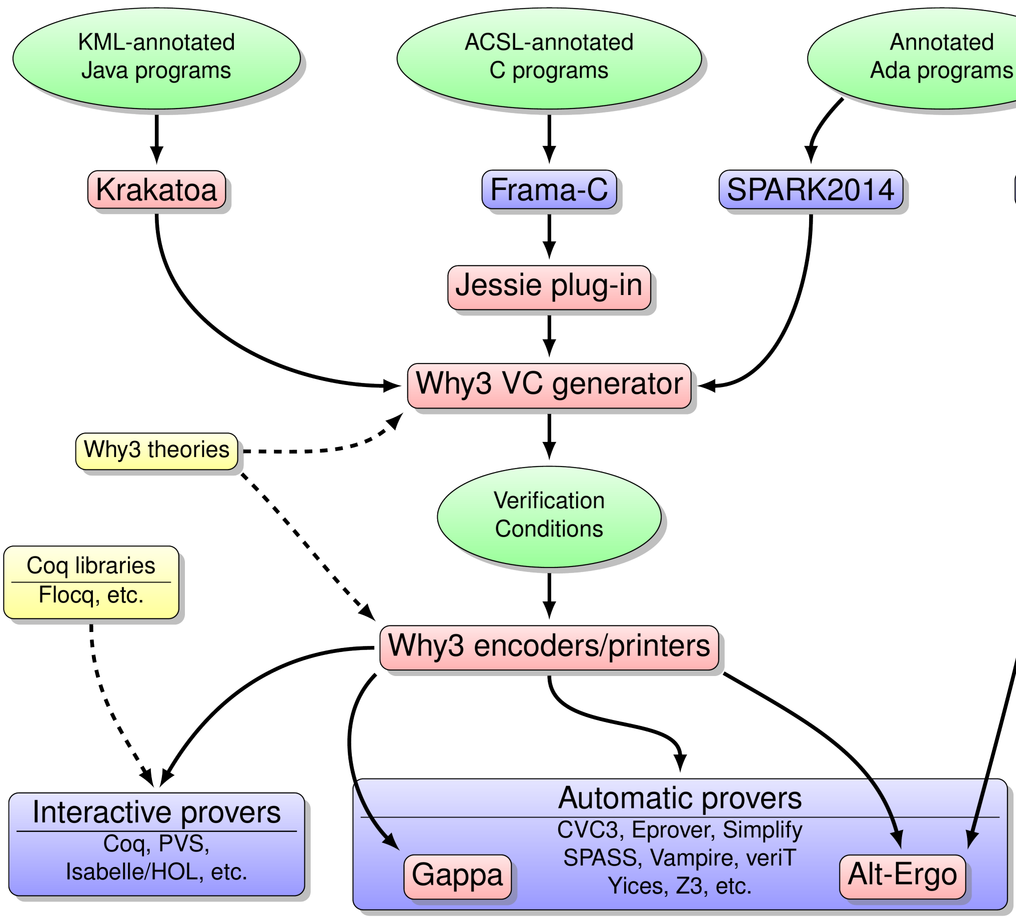


Figure 1. The Why3 ecosystem

3.2.1. The Why3 Ecosystem

This ecosystem is central in our work; it is displayed on Figure 1. The boxes in red background correspond to the tools we develop in the Toccata team.

- The initial design of Why3 was presented in 2012 [55], [99]. In the past years, the main improvements concern the specification language (such as support for higher-order logic functions [76]) and the support for provers. Several new interactive provers are now supported: PVS 6 (used at NASA), Isabelle2014 (planned to be used in the context of Ada program via Spark), and Mathematica. We also added support for new automated provers: CVC4, Metitarski, Metis, Beagle, Princess, and Yices2. More technical improvements are the design of a Coq tactic to call provers via Why3 from Coq, and the design of a proof session mechanism [54]. Why3 was presented during several invited talks [98], [97], [94], [95].
- At the level of the C front-end of Why3 (via Frama-C), we have proposed an approach to add a notion of refinement on C programs [127], and an approach to reason about pointer programs with a standard logic, via *separation predicates* [53]
- The Ada front-end of Why3 has mainly been developed during the past three years, leading to the release of SPARK2014 [105] (http://www.spark-2014.org/)
- In collaboration with J. Almeida, M. Barbosa, J. Pinto, and B. Vieira (University do Minho, Braga, Portugal), J.-C. Filliâtre has developed a method for certifying programs involving cryptographic methods. It uses Why as an intermediate language [45].
- With M. Pereira and S. Melo de Sousa (Universidade da Beira Interior, Covilhã, Portugal), J.-C. Filliâtre has developed an environment for proving ARM assembly code. It uses Why3 as an intermediate VC generator. It was presented at the Inforum conference [124] (best student paper).

3.2.2. Concurrent Programming

- S. Conchon and A. Mebsout, in collaboration with F. Zaïdi (VALS team, LRI), A. Goel and S. Krstić (Strategic Cad Labs, INTEL) have proposed a new model-checking approach for verifying safety properties of array-based systems. This is a syntactically restricted class of parametrized transition systems with states represented as arrays indexed by an arbitrary number of processes. Cache coherence protocols and mutual exclusion algorithms are typical examples of such systems. It was first presented at CAV 2012 [5] and detailed further [87]. It was applied to the verification of programs with fences [83]. The core algorithm has been extended with a mechanism for inferring invariants. This new algorithm, called BRAB, is able to automatically infer invariants strong enough to prove industrial cache coherence protocols. BRAB computes over-approximations of backward reachable states that are checked to be unreachable in a finite instance of the system. These approximations (candidate invariants) are then model-checked together with the original safety properties. Completeness of the approach is ensured by a mechanism for backtracking on spurious traces introduced by too coarse approximations [84], [116].
- In the context of the ERC DeepSea project [0], A. Charguéraud and his co-authors have developed a unifying semantics for various different paradigms of parallel computing (fork-join, async-finish, and futures), and published a conference paper describing this work [44]. Besides, A. Charguéraud and his co-authors have polished their previous work on granularity control for parallel algorithms using user-provided complexity functions, and produced a journal article [43].

3.2.3. Case Studies

- To provide an easy access to the case studies that we develop using Why3 and its front-ends, we have published a *gallery of verified programs* on our web page http://toccata.lri.fr/gallery/. Part of these examples are the solutions to the competitions VerifyThis 2011 [70], VerifyThis 2012 [2], and the competition VScomp 2011 [100].
- Other case studies that led to publications are the design of a library of data-structures based on AVLs [75], and the verification a two-lines C program (solving the N-queens puzzle) using Why3 [96].

[0] Arthur Charguéraud is involved 40% of his time in the ERC DeepSea project, which is hosted at Inria Paris Rocquencourt (team Gallium).

- A. Charguéraud, with F. Pottier (Inria Paris), extended their formalization of the correctness and asympotic complexity of the classic Union Find data structure, which features the bound expressed in terms of the inverse Ackermann function [42]. The proof, conducted using CFML extended with time credits, was refined using a slightly more complex potential function, allowing to derive a simpler and richer interface for the data structure [73].

For other case studies, see also sections of numerical programs and formalization of languages and tools.

3.2.4. Project-team Positioning

Several research groups in the world develop their own approaches, techniques, and tools for deductive verification. With respect to all these related approaches and tools, our originality is our will to use more sophisticated specification languages (with inductive definitions, higher-order features and such) and the ability to use a large set of various theorem provers, including the use of interactive theorem proving to deal with complex functional properties.

- The RiSE team [0] at Microsoft Research Redmond, USA, partly in collaboration with team "programming methodology" team [0] at ETH Zurich develop tools that are closely related to ours: Boogie and Dafny are direct competitors of Why3, VCC is a direct competitor of Frama-C/Jessie.
- The KeY project [0] (several teams, mainly at Karlsruhe and Darmstadt, Germany, and Göteborg, Sweden) develops the KeY tool for Java program verification [41], based on dynamic logic, and has several industrial users. They use a specific modal logic (dynamic logic) for modeling programs, whereas we use standard logic, so as to be able to use off-the-shelf automated provers.
- The "software engineering" group at Augsburg, Germany, develops the KIV system [0], which was created more than 20 years ago (1992) and is still well maintained and efficient. It provides a semi-interactive proof environment based on algebraic-style specifications, and is able to deal with several kinds of imperative style programs. They have a significant industrial impact.
- The VeriFast system [0] aims at verifying C programs specified in Separation Logic. It is developed at the Katholic University at Leuven, Belgium. We do not usually use separation logic (so as to use off-the-shelf provers) but alternative approaches (e.g. static memory separation analysis).
- The Mobius Program Verification Environment [0] is a joint effort for the verification of Java source annotated with JML, combining static analysis and runtime checking. The tool ESC/Java2 [0] is a VC generator similar to Krakatoa, that builds on top of Boogie. It is developed by a community leaded by University of Copenhagen, Denmark. Again, our specificity with respect to them is the consideration of more complex specification languages and interactive theorem proving.
- The Lab for Automated Reasoning and Analysis [0] at EPFL, develop methods and tools for verification of Java (Jahob) and Scala (Leon) programs. They share with us the will and the ability to use several provers at the same time.
- The TLA environment [0], developed by Microsoft Research and the Inria team Veridis, aims at the verification of concurrent programs using mathematical specifications, model checking, and interactive or automated theorem proving.
- The F* project [0], developed by Microsoft Research and the Inria Prosecco team, aims at providing a rich environment for developing programs and proving them.

[0] http://research.microsoft.com/en-us/groups/rise/default.aspx
[0] http://www.pm.inf.ethz.ch/
[0] http://www.key-project.org/
[0] http://www.isse.uni-augsburg.de/en/software/kiv/
[0] http://people.cs.kuleuven.be/~bart.jacobs/verifast/
[0] http://kindsoftware.com/products/opensource/Mobius/
[0] http://kindsoftware.com/products/opensource/ESCJava2/
[0] http://lara.epfl.ch/w/
[0] http://research.microsoft.com/en-us/um/people/lamport/tla/tla.html
[0] http://research.microsoft.com/en-us/projects/fstar/

The KeY and KIV environments mentioned above are partly based on interactive theorem provers. There are other approaches on top of general-purpose proof assistants for proving programs that are not purely functional:

- The Ynot project [0] is a Coq library for writing imperative programs specified in separation logic. It was developed at Harvard University, until the end of the project in 2010. Ynot had similar goals as CFML, although Ynot requires programs to be written in monadic style inside Coq, whereas CFML applies directly on programs written in OCaml syntax, translating them into logical formulae.
- Front-ends to Isabelle were developed to deal with simple sequential imperative programs [126] or C programs [123]. The L4-verified project [106] is built on top of Isabelle.

3.3. Automated Reasoning

Permanent researchers: S. Conchon, G. Melquiond, A. Paskevich

3.3.1. Generalities on Automated Reasoning

- J. C. Blanchette and A. Paskevich have designed an extension to the TPTP TFF (Typed First-order Form) format of theorem proving problems to support rank-1 polymorphic types (also known as ML-style parametric polymorphism) [51]. This extension, named TFF1, has been incorporated in the TPTP standard.
- S. Conchon defended his *habilitation à diriger des recherches* in December 2012. The memoir [80] provides a useful survey of the scientific work of the past 10 years, around the SMT solving techniques, that led to the tools Alt-Ergo and Cubicle as they are nowadays.

3.3.2. Quantifiers and Triggers

- C. Dross, J. Kanig, S. Conchon, and A. Paskevich have proposed a generic framework for adding a decision procedure for a theory or a combination of theories to an SMT prover. This mechanism is based on the notion of instantiation patterns, or *triggers*, which restrict instantiation of universal premises and can effectively prevent a combinatorial explosion. A user provides an axiomatization with triggers, along with a proof of completeness and termination in the proposed framework, and obtains in return a sound, complete and terminating solver for his theory. A prototype implementation was realized on top of Alt-Ergo. As a case study, a feature-rich axiomatization of doubly-linked lists was proved complete and terminating [92]. C. Dross defended her PhD thesis in April 2014 [93]. The main results of the thesis are: (1) a formal semantics of the notion of *triggers* typically used to control quantifier instantiation in SMT solvers, (2) a general setting to show how a first-order axiomatization with triggers can be proved correct, complete, and terminating, and (3) an extended DPLL(T) algorithm to integrate a first-order axiomatization with triggers as a decision procedure for the theory it defines. Significant case studies were conducted on examples coming from SPARK programs, and on the benchmarks on B set theory constructed within the BWare project.

3.3.3. Reasoning Modulo Theories

- S. Conchon, É. Contejean and M. Iguernelala have presented a modular extension of ground AC-completion for deciding formulas in the combination of the theory of equality with user-defined AC symbols, uninterpreted symbols and an arbitrary signature-disjoint Shostak theory X [82]. This work extends the results presented in [81] by showing that a simple preprocessing step allows to get rid of a full AC-compatible reduction ordering, and to simply use a partial multiset extension of a *non-necessarily AC-compatible* ordering.
- S. Conchon, M. Iguernelala, and A. Mebsout have designed a collaborative framework for reasoning modulo simple properties of non-linear arithmetic [86]. This framework has been implemented in the Alt-Ergo SMT solver.

[0] http://ynot.cs.harvard.edu/

- S. Conchon, G. Melquiond and C. Roux have described a dedicated procedure for a theory of floating-point numbers which allows reasoning on approximation errors. This procedure is based on the approach of the Gappa tool: it performs saturation of consequences of the axioms, in order to refine bounds on expressions. In addition to the original approach, bounds are further refined by a constraint solver for linear arithmetic [88]. This procedure has been implemented in Alt-Ergo.
- In collaboration with A. Mahboubi (Inria project-team Typical), and G. Melquiond, the group involved in the development of Alt-Ergo have implemented and proved the correctness of a novel decision procedure for quantifier-free linear integer arithmetic [1]. This algorithm tries to bridge the gap between projection and branching/cutting methods: it interleaves an exhaustive search for a model with bounds inference. These bounds are computed provided an oracle capable of finding constant positive linear combinations of affine forms. An efficient oracle based on the Simplex procedure has been designed. This algorithm is proved sound, complete, and terminating and is implemented in Alt-Ergo.
- Most of the results above are detailed in M. Iguernelala's PhD thesis [103].

3.3.4. Applications

- We have been quite successful in the application of Alt-Ergo to industrial development: qualification by Airbus France, integration of Alt-Ergo into the Spark Pro toolset.
- In the context of the BWare project, aiming at using Why3 and Alt-Ergo for discharging proof obligations generated by Atelier B, we made progress into several directions. The method of translation of B proof obligations into Why3 goals was first presented at ABZ'2012 [119]. Then, new drivers have been designed for Why3, in order to use new back-end provers Zenon modulo and iProver modulo. A notion of rewrite rule was introduced into Why3, and a transformation for simplifying goals before sending them to back-end provers was designed. Intermediate results obtained so far in the project were presented both at the French conference AFADL [91] and at ABZ'2014 [90].

  On the side of Alt-Ergo, recent developments have been made to efficiently discharge proof obligations generated by Atelier B. This includes a new plugin architecture to facilitate experiments with different SAT engines, new heuristics to handle quantified formulas, and important modifications in its internal data structures to boost performances of core decision procedures. Benchmarks realized on more than 10,000 proof obligations generated from industrial B projects show significant improvements [85].
- Hybrid automatons interleave continuous behaviors (described by differential equations) with discrete transitions. D. Ishii and G. Melquiond have worked on an automated procedure for verifying safety properties (that is, global invariants) of such systems [104].

3.3.5. Project-team Positioning

Automated Theorem Proving is a large community, but several sub-groups can be identified:

- The SMT-LIB community gathers people interested in reasoning modulo theories. In this community, only a minority of participants are interested in supporting first-order quantifiers at the same time as theories. SMT solvers that support quantifiers are Z3 (Microsoft Research Redmond, USA), CVC3 and its successor CVC4 [0].
- The TPTP community gathers people interested in first-order theorem proving.
- Other Inria teams develop provers: veriT by team Veridis, and Psyche by team Parsifal.
- Other groups develop provers dedicated to very specific cases, such as Metitarski [0] at Cambridge, UK, which aims at proving formulas on real numbers, in particular involving special functions such as log or exp. The goal is somewhat similar to our CoqInterval library, *cf* objective 4.

[0] http://cvc4.cs.nyu.edu/web/
[0] http://www.cl.cam.ac.uk/~lp15/papers/Arith/

It should be noticed that a large number of provers mentioned above are connected to Why3 as back-ends.

3.4. Formalization and Certification of Languages, Tools and Systems

Permanent researchers: S. Boldo, A. Charguéraud, C. Marché, G. Melquiond, C. Paulin

3.4.1. Real Numbers, Real Analysis, Probabilities

- S. Boldo, C. Lelay, and G. Melquiond have worked on the Coquelicot library, designed to be a user-friendly Coq library about real analysis [65], [66]. An easier way of writing formulas and theorem statements is achieved by relying on total functions in place of dependent types for limits, derivatives, integrals, power series, and so on. To help with the proof process, the library comes with a comprehensive set of theorems and some automation. We have exercised the library on several use cases: on an exam at university entry level [108], for the definitions and properties of Bessel functions [107], and for the solution of the one-dimensional wave equation [109]. We have also conducted a survey on the formalization of real arithmetic and real analysis in various proof systems [67].
- Watermarking techniques are used to help identify copies of publicly released information. They consist in applying a slight and secret modification to the data before its release, in a way that should remain recognizable even in (reasonably) modified copies of the data. Using the Coq ALEA library, which formalizes probability theory and probabilistic programs, D. Baelde together with P. Courtieu, D. Gross-Amblard from Rennes and C. Paulin have established new results about the robustness of watermarking schemes against arbitrary attackers [47]. The technique for proving robustness is adapted from methods commonly used for cryptographic protocols and our work illustrates the strengths and particularities of the ALEA style of reasoning about probabilistic programs.

3.4.2. Formalization of Languages, Semantics

- P. Herms, together with C. Marché and B. Monate (CEA List), has developed a certified VC generator, using Coq. The program for VC calculus and its specifications are both written in Coq, but the code is crafted so that it can be extracted automatically into a stand-alone executable. It is also designed in a way that allows the use of arbitrary first-order theorem provers to discharge the generated obligations [102]. On top of this generic VC generator, P. Herms developed a certified VC generator for C source code annotated using ACSL. This work is the main result of his PhD thesis [101].
- A. Tafat and C. Marché have developed a certified VC generator using Why3 [112], [113]. The challenge was to formalize the operational semantics of an imperative language, and a corresponding weakest precondition calculus, without the possibility to use Coq advanced features such as dependent types or higher-order functions. The classical issues with local bindings, names and substitutions were solved by identifying appropriate lemmas. It was shown that Why3 can offer a significantly higher amount of proof automation compared to Coq.
- A. Charguéraud, together with Alan Schmitt (Inria Rennes) and Thomas Wood (Imperial College), has developed an interactive debugger for JavaScript. The interface, accessible as a webpage in a browser, allows to execute a given JavaScript program, following step by step the formal specification of JavaScript developped in prior work on *JsCert* [56]. Concretely, the tool acts as a double-debugger: one can visualize both the state of the interpreted program and the state of the interpreter program. This tool is intended for the JavaScript committee, VM developpers, and other experts in JavaScript semantics.
- M. Clochard, C. Marché, and A. Paskevich have developed a general setting for developing programs involving binders, using Why3. This approach was successfully validated on two case studies: a verified implementation of untyped lambda-calculus and a verified tableaux-based theorem prover [79].

- M. Clochard, J.-C. Filliâtre, C. Marché, and A. Paskevich have developed a case study on the formalization of semantics of programming languages using Why3 [76]. This case study aims at illustrating recent improvements of Why3 regarding the support for higher-order logic features in the input logic of Why3, and how these are encoded into first-order logic, so that goals can be discharged by automated provers. This case study also illustrates how reasoning by induction can be done without need for interactive proofs, via the use of *lemma functions*.
- M. Clochard and L. Gondelman have developed a formalization of a simple compiler in Why3 [77]. It compiles a simple imperative language into assembler instructions for a stack machine. This case study was inspired by a similar example developed using Coq and interactive theorem proving. The aim is to improve significantly the degree of automation in the proofs. This is achieved by the formalization of a Hoare logic and a Weakest Precondition Calculus on assembly programs, so that the correctness of compilation is seen as a formal specification of the assembly instructions generated.

3.4.3. Project-team Positioning

The objective of formalizing languages and algorithms is very general, and it is pursued by several Inria teams. One common trait is the use of the Coq proof assistant for this purpose: Pi.r2 (development of Coq itself and its meta-theory), Gallium (semantics and compilers of programming languages), Marelle (formalization of mathematics), SpecFun (real arithmetic), Celtique (formalization of static analyzers).

Other environments for the formalization of languages include

- ACL2 system [0]: an environment for writing programs with formal specifications in first-order logic based on a Lisp engine. The proofs are conducted using a prover based on the Boyer-Moore approach. It is a rather old system but still actively maintained and powerful, developed at University of Texas at Austin. It has a strong industrial impact.
- Isabelle environment [0]: both a proof assistant and an environment for developing pure applicative programs. It is developed jointly at University of Cambridge, UK, Technische Universität München, Germany, and to some extent by the VALS team at LRI, Université Paris-Sud. It features highly automated tactics based on ATP systems (the Sledgehammer tool).
- The team “Trustworthy Systems” at NICTA in Australia [0] aims at developing highly trustable software applications. They developed a formally verified micro-kernel called seL4 [106], using a home-made layer to deal with C programs on top of the Isabelle prover.
- The PVS system [0] is an environment for both programming and proving (purely applicative) programs. It is developed at the Computer Science Laboratory of SRI international, California, USA. A major user of PVS is the team LFM [0] at NASA Langley, USA, for the certification of programs related to air traffic control.

In the Toccata team, we do not see these alternative environments as competitors, even though, for historical reasons, we are mainly using Coq. Indeed both Isabelle and PVS are available as back-ends of Why3.

3.5. Proof of Numerical Programs

Permanent researchers: S. Boldo, C. Marché, G. Melquiond

- Linked with objective 1 (Deductive Program Verification), the methodology for proving numerical C programs has been presented by S. Boldo in her habilitation [58] and as invited speaker [59]. An application is the formal verification of a numerical analysis program. S. Boldo, J.-C. Filliâtre, and G. Melquiond, with F. Clément and P. Weis (POMDAPI team, Inria Paris - Rocquencourt), and M. Mayero (LIPN), completed the formal proof of the second-order centered finite-difference scheme for the one-dimensional acoustic wave [61][3].

[0] http://www.cs.utexas.edu/~moore/acl2/
[0] http://isabelle.in.tum.de/
[0] http://ssrg.nicta.com.au/projects/TS/
[0] http://pvs.csl.sri.com/
[0] http://shemesh.larc.nasa.gov/fm/fm-main-team.html

- Several challenging floating-point algorithms have been studied and proved. This includes an algorithm by Kahan for computing the area of a triangle: S. Boldo proved an improvement of its error bound and new investigations in case of underflow [57]. This includes investigations about quaternions. They should be of norm 1, but due to the round-off errors, a drift of this norm is observed over time. C. Marché determined a bound on this drift and formally proved it correct [8]. P. Roux formally verified an algorithm for checking that a matrix is semi-definite positive [125]. The challenge here is that testing semi-definiteness involves algebraic number computations, yet it needs to be implemented using only approximate floating-point operations.
- Because of compiler optimizations (or bugs), the floating-point semantics of a program might change once compiled, thus invalidating any property proved on the source code. We have investigated two ways to circumvent this issue, depending on whether the compiler is a black box. When it is, T. Nguyen has proposed to analyze the assembly code it generates and to verify it is correct [122]. On the contrary, S. Boldo and G. Melquiond (in collaboration with J.-H. Jourdan and X. Leroy) have added support for floating-point arithmetic to the CompCert compiler and formally proved that none of the transformations the compiler applies modify the floating-point semantics of the program [64], [63].
- Linked with objectives 2 (Automated Reasoning) and 3 (Formalization and Certification of Languages, Tools and Systems), G. Melquiond has implemented an efficient Coq library for floating-point arithmetic and proved its correctness in terms of operations on real numbers [117]. It serves as a basis for an interval arithmetic on which Taylor models have been formalized. É. Martin-Dorel and G. Melquiond have integrated these models into CoqInterval [9]. This Coq library is dedicated to automatically proving the approximation properties that occur when formally verifying the implementation of mathematical libraries (libm).
- Double rounding occurs when the target precision of a floating-point computation is narrower than the working precision. In some situations, this phenomenon incurs a loss of accuracy. P. Roux has formally studied when it is innocuous for basic arithmetic operations [125]. É. Martin-Dorel and G. Melquiond (in collaboration with J.-M. Muller) have formally studied how it impacts algorithms used for error-free transformations [115]. These works were based on the Flocq formalization of floating-point arithmetic for Coq.
- By combining multi-precision arithmetic, interval arithmetic, and massively-parallel computations, G. Melquiond (in collaboration with G. Nowak and P. Zimmermann) has computed enough digits of the Masser-Gramain constant to invalidate a 30-year old conjecture about its closed form [118].

3.5.1. Project-team Positioning

This objective deals both with formal verification and floating-point arithmetic, which is quite uncommon. Therefore our competitors/peers are few. We may only cite the works by J. Duracz and M. Konečný, Aston University in Birmingham, UK.

The Inria team AriC (Grenoble - Rhône-Alpes) is closer to our research interests, but they are lacking manpower on the formal proof side; we have numerous collaborations with them. The Inria team Caramel (Nancy - Grand Est) also shares some research interests with us, though fewer; again, they do not work on the formal aspect of the verification; we have some occasional collaborations with them.

There are many formalization efforts from chip manufacturers, such as AMD (using the ACL2 proof assistant) and Intel (using the Forte proof assistants) but the algorithms they consider are quite different from the ones we study. The works on the topic of floating-point arithmetic from J. Harrison at Intel using HOL Light are really close to our research interests, but they seem to be discontinued.

A few deductive program verification teams are willing to extend their tools toward floating-point programs. This includes the KeY project and SPARK. We have an ongoing collaboration with the latter, in the context of the ProofInUSe project.

Deductive verification is not the only way to prove programs. Abstract interpretation is widely used, and several teams are interested in floating-point arithmetic. This includes the Inria team Antique (Paris - Rocquencourt) and a CEA List team, who have respectively developed the Astrée and Fluctuat tools. This approach targets a different class of numerical algorithms than the ones we are interested in.

Other people, especially from the SMT community (*cf* objective 2), are also interested in automatically proving formulas about floating-point numbers, notably at Oxford University. They are mainly focusing on pure floating-point arithmetic though and do not consider them as approximation of real numbers.

Finally, it can be noted that numerous teams are working on the verification of numerical programs, but assuming the computations are real rather than floating-point ones. This is out of the scope of this objective.

4. Application Domains

4.1. Domain

The application domains we target involve safety-critical software, that is where a high-level guarantee of soundness of functional execution of the software is wanted. Currently our industrial collaborations mainly belong to the domain of transportation, including aeronautics, railroad, space flight, automotive.

Verification of C programs, Alt-Ergo at Airbus Transportation is the domain considered in the context of the ANR U3CAT project, led by CEA, in partnership with Airbus France, Dassault Aviation, Sagem Défense et Sécurité. It included proof of C programs via Frama-C/Jessie/Why, proof of floating-point programs [114], the use of the Alt-Ergo prover via CAVEAT tool (CEA) or Frama-C/WP. Within this context, we contributed to a qualification process of Alt-Ergo with Airbus industry: the technical documents (functional specifications and benchmark suite) have been accepted by Airbus, and these documents were submitted by Airbus to the certification authorities (DO-178B standard) in 2012. This action is continued in the new project Soprano.

Certified compilation, certified static analyzers Aeronautics is the main target of the Verasco project, led by Verimag, on the development of certified static analyzers, in partnership with Airbus. This is a follow-up of the transfer of the CompCert certified compiler (Inria team Gallium) to which we contributed to the support of floating-point computations [64].

Transfer to the community of Ada development The former FUI project Hi-Lite, led by Adacore company, introduced the use of Why3 and Alt-Ergo as back-end to SPARK2014, an environment for verification of Ada programs. This is applied to the domain of aerospace (Thales, EADS Astrium). At the very beginning of that project, Alt-Ergo was added in the Spark Pro toolset (predecessor of SPARK2014), developed by Altran-Praxis: Alt-Ergo can be used by customers as an alternate prover for automatically proving verification conditions. Its usage is described in the new edition of the Spark book [0] (Chapter "Advanced proof tools"). This action is continued in the new joint laboratory ProofInUse. A recent paper [72] provides an extensive list of applications of SPARK, a major one being the British air control management *iFacts*.

Transfer to the community of Atelier B In the current ANR project BWare, we investigate the use of Why3 and Alt-Ergo as an alternative back-end for checking proof obligations generated by *Atelier B*, whose main applications are railroad-related software,

a collaboration with Mitsubishi Electric R&D Centre Europe (Rennes) (joint publication [119]) and ClearSy (Aix-en-Provence).

SMT-based Model-Checking: Cubicle S. Conchon (with A. Mebsout and F. Zaidi from VALS team at LRI) has a long-term collaboration with S. Krstic and A. Goel (Intel Strategic Cad Labs in Hillsboro, OR, USA) that aims in the development of the SMT-based model checker Cubicle (http://cubicle.lri.fr/) based on Alt-Ergo [116][5]. It is particularly targeted to the verification of concurrent programs and protocols.

[0] http://www.altran-praxis.com/book/

5. Highlights of the Year

5.1. Highlights of the Year

S. Conchon has co-organized POPL'2017 (January, Paris, http://conf.researchr.org/home/POPL-2017).

C. Marché has co-organized the first joint Frama-C/SPARK day (May, Paris, http://frama-c.com/FCSD17.html), in the context of the Open Source Innovation Spring (http://www.open-source-innovation-spring.org/).

S. Boldo and G. Melquiond have published a book: Computer Arithmetic and Formal Proofs, Verifying Floating-point Algorithms with the Coq System [32].

5.1.1. Awards

M. Pereira and R. Rieu-Helft received the "Best student team" award, and J.-C. Filliâtre the "Best overall team" award, at the *VerifyThis@ETAPS2017 verification competition*.

6. New Software and Platforms

6.1. Alt-Ergo

Automated theorem prover for software verification
KEYWORDS: Software Verification - Automated theorem proving
FUNCTIONAL DESCRIPTION: Alt-Ergo is an automatic solver of formulas based on SMT technology. It is especially designed to prove mathematical formulas generated by program verification tools, such as Frama-C for C programs, or SPARK for Ada code. Initially developed in Toccata research team, Alt-Ergo's distribution and support are provided by OCamlPro since September 2013.

RELEASE FUNCTIONAL DESCRIPTION: the "SAT solving" part can now be delegated to an external plugin, new experimental SAT solver based on mini-SAT, provided as a plugin. This solver is, in general, more efficient on ground problems, heuristics simplification in the default SAT solver and in the matching (instantiation) module, re-implementation of internal literals representation, improvement of theories combination architecture, rewriting some parts of the formulas module, bugfixes in records and numbers modules, new option "-no-Ematching" to perform matching without equality reasoning (i.e. without considering "equivalence classes"). This option is very useful for benchmarks coming from Atelier-B, two new experimental options: "-save-used-context" and "-replay-used-context". When the goal is proved valid, the first option allows to save the names of useful axioms into a ".used" file. The second one is used to replay the proof using only the axioms listed in the corresponding ".used" file. Note that the replay may fail because of the absence of necessary ground terms generated by useless axioms (that are not included in .used file) during the initial run.

- Participants: Alain Mebsout, Évelyne Contejean, Mohamed Iguernelala, Stéphane Lescuyer and Sylvain Conchon
- Partner: OCamlPro
- Contact: Sylvain Conchon
- URL: http://alt-ergo.lri.fr

6.2. CFML

Interactive program verification using characteristic formulae
KEYWORDS: Coq - Software Verification - Deductive program verification - Separation Logic

FUNCTIONAL DESCRIPTION: The CFML tool supports the verification of OCaml programs through interactive Coq proofs. CFML proofs establish the full functional correctness of the code with respect to a specification. They may also be used to formally establish bounds on the asymptotic complexity of the code. The tool is made of two parts: on the one hand, a characteristic formula generator implemented as an OCaml program that parses OCaml code and produces Coq formulae, and, on the other hand, a Coq library that provides notations and tactics for manipulating characteristic formulae interactively in Coq.

- Participants: Arthur Charguéraud, Armaël Guéneau and François Pottier
- Contact: Arthur Charguéraud
- URL: http://www.chargueraud.org/softs/cfml/

6.3. Coq

The Coq Proof Assistant
KEYWORDS: Proof - Certification - Formalisation
SCIENTIFIC DESCRIPTION: Coq is an interactive proof assistant based on the Calculus of (Co-)Inductive Constructions, extended with universe polymorphism. This type theory features inductive and co-inductive families, an impredicative sort and a hierarchy of predicative universes, making it a very expressive logic. The calculus allows to formalize both general mathematics and computer programs, ranging from theories of finite structures to abstract algebra and categories to programming language metatheory and compiler verification. Coq is organised as a (relatively small) kernel including efficient conversion tests on which are built a set of higher-level layers: a powerful proof engine and unification algorithm, various tactics/decision procedures, a transactional document model and, at the very top an IDE.
FUNCTIONAL DESCRIPTION: Coq provides both a dependently-typed functional programming language and a logical formalism, which, altogether, support the formalisation of mathematical theories and the specification and certification of properties of programs. Coq also provides a large and extensible set of automatic or semi-automatic proof methods. Coq's programs are extractible to OCaml, Haskell, Scheme, ...

RELEASE FUNCTIONAL DESCRIPTION: Version 8.7 features a large amount of work on cleaning and speeding up the code base, notably the work of Pierre-Marie Pédrot on making the tactic-level system insensitive to existential variable expansion, providing a safer API to plugin writers and making the code more robust.

New tactics: Variants of tactics supporting existential variables "eassert", "eenough", etc. by Hugo Herbelin. Tactics "extensionality in H" and "inversion_sigma" by Jason Gross, "specialize with" accepting partial bindings by Pierre Courtieu.

Cumulative Polymorphic Inductive Types, allowing cumulativity of universes to go through applied inductive types, by Amin Timany and Matthieu Sozeau.

The SSReflect plugin by Georges Gonthier, Assia Mahboubi and Enrico Tassi was integrated (with its documentation in the reference manual) by Maxime Dénès, Assia Mahboubi and Enrico Tassi.

The "coq_makefile" tool was completely redesigned to improve its maintainability and the extensibility of generated Makefiles, and to make "_CoqProject" files more palatable to IDEs by Enrico Tassi.

A lot of other changes are described in the CHANGES file.
NEWS OF THE YEAR: Version 8.7 was released in October 2017 and version 8.7.1 in December 2017, development started in January 2017. This is the second release of Coq developed on a time-based development cycle. Its development spanned 9 months from the release of Coq 8.6 and was based on a public road-map. It attracted many external contributions. Code reviews and continuous integration testing were systematically used before integration of new features, with an important focus given to compatibility and performance issues.

The main scientific advance in this version is the integration of cumulative inductive types in the system. More practical advances in stability, performance, usability and expressivity of tactics were also implemented, resulting in a mostly backwards-compatible but appreciably faster and more robust release. Much work on plugin extensions to Coq by the same development team has also been going on in parallel, including work on JSCoq by Emilio JG Arias, Ltac 2 by P.M-Pédrot, which required synchronised changes of the main codebase. In 2017, the construction of the Coq Consortium by Yves Bertot and Maxime Dénès has greatly advanced and is now nearing its completion.

- Participants: Abhishek Anand, C. J. Bell, Yves Bertot, Frédéric Besson, Tej Chajed, Pierre Courtieu, Maxime Denes, Julien Forest, Emilio Jesús Gallego Arias, Gaëtan Gilbert, Benjamin Grégoire, Jason Gross, Hugo Herbelin, Ralf Jung, Matej Kosik, Sam Pablo Kuper, Xavier Leroy, Pierre Letouzey, Assia Mahboubi, Cyprien Mangin, Érik Martin-Dorel, Olivier Marty, Guillaume Melquiond, Pierre-Marie Pédrot, Benjamin C. Pierce, Lars Rasmusson, Yann Régis-Gianas, Lionel Rieg, Valentin Robert, Thomas Sibut-Pinote, Michael Soegtrop, Matthieu Sozeau, Arnaud Spiwack, Paul Steckler, George Stelle, Pierre-Yves Strub, Enrico Tassi, Hendrik Tews, Laurent Théry, Amin Timany, Vadim Zaliva and Théo Zimmermann
- Partners: CNRS - Université Paris-Sud - ENS Lyon - Université Paris-Diderot
- Contact: Matthieu Sozeau
- Publication: The Coq Proof Assistant, version 8.7.1
- URL: http://coq.inria.fr/

6.4. CoqInterval

Interval package for Coq
KEYWORDS: Interval arithmetic - Coq
FUNCTIONAL DESCRIPTION: CoqInterval is a library for the proof assistant Coq.

It provides several tactics for proving theorems on enclosures of real-valued expressions. The proofs are performed by an interval kernel which relies on a computable formalization of floating-point arithmetic in Coq.

The Marelle team developed a formalization of rigorous polynomial approximation using Taylor models in Coq. In 2014, this library has been included in CoqInterval.

- Participants: Assia Mahboubi, Érik Martin-Dorel, Guillaume Melquiond, Jean-Michel Muller, Laurence Rideau, Laurent Théry, Micaela Mayero, Mioara Joldes, Nicolas Brisebarre and Thomas Sibut-Pinote
- Contact: Guillaume Melquiond
- Publications: Proving bounds on real-valued functions with computations - Floating-point arithmetic in the Coq system - Proving Tight Bounds on Univariate Expressions with Elementary Functions in Coq - Formally Verified Approximations of Definite Integrals - Formally Verified Approximations of Definite Integrals
- URL: http://coq-interval.gforge.inria.fr/

6.5. Coquelicot

The Coquelicot library for real analysis in Coq
KEYWORDS: Coq - Real analysis

FUNCTIONAL DESCRIPTION: Coquelicot is library aimed for supporting real analysis in the Coq proof assistant. It is designed with three principles in mind. The first is the user-friendliness, achieved by implementing methods of automation, but also by avoiding dependent types in order to ease the stating and readability of theorems. This latter part was achieved by defining total function for basic operators, such as limits or integrals. The second principle is the comprehensiveness of the library. By experimenting on several applications, we ensured that the available theorems are enough to cover most cases. We also wanted to be able to extend our library towards more generic settings, such as complex analysis or Euclidean spaces. The third principle is for the Coquelicot library to be a conservative extension of the Coq standard library, so that it can be easily combined with existing developments based on the standard library.

- Participants: Catherine Lelay, Guillaume Melquiond and Sylvie Boldo
- Contact: Sylvie Boldo
- URL: http://coquelicot.saclay.inria.fr/

6.6. Cubicle

The Cubicle model checker modulo theories
KEYWORDS: Model Checking - Software Verification
FUNCTIONAL DESCRIPTION: Cubicle is an open source model checker for verifying safety properties of array-based systems, which corresponds to a syntactically restricted class of parametrized transition systems with states represented as arrays indexed by an arbitrary number of processes. Cache coherence protocols and mutual exclusion algorithms are typical examples of such systems.

- Participants: Alain Mebsout and Sylvain Conchon
- Contact: Sylvain Conchon
- URL: http://cubicle.lri.fr/

6.7. Flocq

The Flocq library for formalizing floating-point arithmetic in Coq
KEYWORDS: Floating-point - Arithmetic code - Coq
FUNCTIONAL DESCRIPTION: The Flocq library for the Coq proof assistant is a comprehensive formalization of floating-point arithmetic: core definitions, axiomatic and computational rounding operations, high-level properties. It provides a framework for developers to formally verify numerical applications.

Flocq is currently used by the CompCert verified compiler to support floating-point computations.

- Participants: Guillaume Melquiond, Pierre Roux and Sylvie Boldo
- Contact: Sylvie Boldo
- Publications: Flocq: A Unified Library for Proving Floating-point Algorithms in Coq - A Formally-Verified C Compiler Supporting Floating-Point Arithmetic - Verified Compilation of Floating-Point Computations - Innocuous Double Rounding of Basic Arithmetic Operations - Formal Proofs of Rounding Error Bounds - Computer Arithmetic and Formal Proofs
- URL: http://flocq.gforge.inria.fr/

6.8. Gappa

The Gappa tool for automated proofs of arithmetic properties
KEYWORDS: Floating-point - Arithmetic code - Software Verification - Constraint solving

FUNCTIONAL DESCRIPTION: Gappa is a tool intended to help formally verifying numerical programs dealing with floating-point or fixed-point arithmetic. It has been used to write robust floating-point filters for CGAL and it is used to verify elementary functions in CRlibm. While Gappa is intended to be used directly, it can also act as a backend prover for the Why3 software verification plateform or as an automatic tactic for the Coq proof assistant.

- Participant: Guillaume Melquiond
- Contact: Guillaume Melquiond
- Publications: Generating formally certified bounds on values and round-off errors - Formal certification of arithmetic filters for geometric predicates - Assisted verification of elementary functions - From interval arithmetic to program verification - Formally Certified Floating-Point Filters For Homogeneous Geometric Predicates - Combining Coq and Gappa for Certifying Floating-Point Programs - Handbook of Floating-Point Arithmetic - Certifying the floating-point implementation of an elementary function using Gappa - Automations for verifying floating-point algorithms - Automating the verification of floating-point algorithms - Computer Arithmetic and Formal Proofs
- URL: http://gappa.gforge.inria.fr/

6.9. Why3

The Why3 environment for deductive verification
KEYWORDS: Formal methods - Trusted software - Software Verification - Deductive program verification
FUNCTIONAL DESCRIPTION: Why3 is an environment for deductive program verification. It provides a rich language for specification and programming, called WhyML, and relies on external theorem provers, both automated and interactive, to discharge verification conditions. Why3 comes with a standard library of logical theories (integer and real arithmetic, Boolean operations, sets and maps, etc.) and basic programming data structures (arrays, queues, hash tables, etc.). A user can write WhyML programs directly and get correct-by-construction OCaml programs through an automated extraction mechanism. WhyML is also used as an intermediate language for the verification of C, Java, or Ada programs.

- Participants: Andriy Paskevych, Claude Marché, François Bobot, Guillaume Melquiond, Jean-Christophe Filliâtre, Levs Gondelmans and Martin Clochard
- Partners: CNRS - Université Paris-Sud
- Contact: Claude Marché
- URL: http://why3.lri.fr/

7. New Results

7.1. Deductive Verification

Synthetic topology in HoTT for probabilistic programming. F. Faissole and B. Spitters have developed a mathematical formalism based on synthetic topology and homotopy type theory to interpret probabilistic algorithms. They suggest to use proof assistants to prove such programs [39] [31]. They also have formalized synthetic topology in the Coq proof assistant using the HoTT library. It consists of a theory of lower reals, valuations and lower integrals. All the results are constructive. They apply their results to interpret probabilistic programs using a monadic approach [28].

Defunctionalization for proving higher-order programs. J.-C. Filliâtre and M. Pereira proposed a new approach to the verification of higher-order programs, using the technique of defunctionalization, that is, the translation of first-class functions into first-order values. This is an early experimental work, conducted on examples only within the Why3 system. This work was published at JFLA 2017 [29].

Extracting Why3 programs to C programs. R. Rieu-Helft, C. Marché, and G. Melquiond devised a simple memory model for representing C-like pointers in the Why3 system. This makes it possible to translate a small fragment of Why3 verified programs into idiomatic C code [30]. This extraction mechanism was used to turn a verified Why3 library of arbitrary-precision integer arithmetic into a C library that can be substituted to part of the GNU Multi-Precision (GMP) library [23].

Verification of highly imperative OCaml programs with Why3 J.-C. Filliâtre, M. Pereira and S. Melo de Sousa proposed a new methodology for proving highly imperative OCaml programs with Why3. For a given OCaml program, a specific memory model is built and one checks a Why3 program that operates on it. Once the proof is complete, they use Why3's extraction mechanism to translate its programs to OCaml, while replacing the operations on the memory model with the corresponding operations on mutable types of OCaml. This method is evaluated on several examples that manipulate linked lists and mutable graphs [20].

7.2. Automated Reasoning

A Three-tier Strategy for Reasoning about Floating-Point Numbers in SMT. The SMT-LIB standard defines a formal semantics for a theory of floating-point (FP) arithmetic (FPA). This formalization reduces FP operations to reals by means of a rounding operator, as done in the IEEE-754 standard. Closely following this description, S. Conchon, M. Iguernlala, K. Ji, G. Melquiond and C. Fumex propose a three-tier strategy to reason about FPA in SMT solvers. The first layer is a purely axiomatic implementation of the automatable semantics of the SMT-LIB standard. It reasons with exceptional cases (e.g. overflows, division by zero, undefined operations) and reduces finite representable FP expressions to reals using the rounding operator. At the core of the strategy, a second layer handles a set of lemmas about the properties of rounding. For these lemmas to be used effectively, the instantiation mechanism of SMT solvers is extended to tightly cooperate with the third layer, the NRA engine of SMT solvers, which provides interval information. The strategy is implemented in the Alt-Ergo SMT solver and validated on a set of benchmarks coming from the SMT-LIB competition, and also from the deductive verification of C and Ada programs. The results show that the approach is promising and compete with existing techniques implemented in state-of-the-art SMT solvers. This work was presented at the CAV conference [18].

Lightweight Approach for Declarative Proofs. M. Clochard designed an extension of first-order logic, for describing reasoning steps needed to discharge a proof obligation. The extension is under the form of two new connectives, called proof indications, that allow the user to encode reasoning steps inside a logic formula. This extension makes possible to use the syntax of formulas as a proof language. The approach was presented at the JFLA conference [26] and implemented in Why3. It brings a lightweight mechanism for declarative proofs in an environment like Why3 where provers are used as black boxes. Moreover, this mechanism restricts the scope of auxiliary lemmas, reducing the size of proof obligations sent to external provers.

7.3. Certification of Algorithms, Languages, Tools and Systems

Formalization and closedness of finite dimensional subspaces. F. Faissole formalized a theory of finite dimensional subspaces of Hilbert spaces in order to apply the Lax-Milgram Theorem on such subspaces. He had to prove, in the Coq proof assistant, that finite dimensional subspaces of Hilbert spaces are closed in the context of general topology using filters [19]. He also formalized both finite dimensional modules and finite dimensional subspaces of modules. He compared the two formalizations and showed a complementarity between them. He proved that the product of two finite dimensional modules is a finite dimensional module [27].

Verified numerical approximations of improper definite integrals. The CoqInterval library provides some tactics for computing and formally verifying numerical approximations of real-valued expressions inside the Coq system. In particular, it is able to compute reliable bounds on proper definite integrals [111]. A. Mahboubi, G. Melquiond, and T. Sibut-Pinote extended these algorithms to also cover some improper integrals, e.g., those with an unbounded integration domain [40]. This makes CoqInterval one of the very few tools able to produce reliable results for improper integrals, be they formally verified or not.

A Coq Formal Proof of the Lax–Milgram theorem. S. Boldo, F. Clément, F. Faissole, V. Martin, and M. Mayero worked on a Coq formal proof of the Lax–Milgram theorem. It is one of the theoretical cornerstone for the correctness of the Finite Element Method. It required many results from linear algebra, geometry, functional analysis, and Hilbert spaces [13] [24].

Formalization of numerical filters S. Boldo, D. Gallois-Wong, and T. Hilaire developped a formalization in the Coq proof assistant of numerical filters. It includes equivalences between several expressions and the formal proof of the Worst-Case Peak Gain Theorem to bound the magnitude of the outputs (and every intern variable) of stable filters.

A Verified OCaml Library. Abstract Libraries are the basic building blocks of any realistic programming project. It is thus of utmost interest for a programmer to build her software on top of bug-free libraries. At the ML family workshop [38], A. Charguéraud, J.-C. Filliâtre, M. Pereira and F. Pottier presented the ongoing VOCAL project, which aims at building a mechanically verified library of general-purpose data structures and algorithms, written in the OCaml language. A key ingredient of VOCAL is the design of a specification language for OCaml, independently of any verification tool.

Formal Analysis of shell scripts. The shell language is widely used for various system administration tasks on UNIX machines. The CoLiS project aims at applying formal methods for verifying scripts used for installation of packages of software distributions. The syntax and semantics of shell are particularly treacherous. They proposed a new language called CoLiS which, on the one hand, has well-defined static semantics and avoids some of the pitfalls of the shell, and, on the other hand, is close enough to the shell to be the target of an automated translation of the scripts in our corpus. In collaboration with N. Jeannerod and R. Treinen, C. Marché formalized the syntax and semantics of CoLiS in Why3, defined an interpreter for the language in the WhyML programming language, and present an automated proof in the Why3 proof environment of soundness and completeness of this interpreter with respect to the formal semantics [22]. The development is available in Toccata's gallery http://toccata.lri.fr/gallery/colis_interpreter.en.html. This formalized interpreter is extracted to OCaml and the verified code is integrated into a prototype software toolset developed by I. Dami and C. Marché [36].

A verified yet efficient arbitrary-precision integer library. R. Rieu-Helft used the Why3 system to implement, specify, and verify a library of arbitrary-precision integer arithmetic: comparison, addition, multiplication, shifts, division. A lot of efforts were put into replicating and verifying the numerous implementation tricks the GMP library uses to achieve state-of-the-art performances, especially for the division algorithm. While the resulting library is nowhere near as fast as the hand-written assembly code GMP uses, it is competitive with the generic C code of GMP for small integers (i.e., mini-GMP) [23]. The development is available in Toccata's gallery http://toccata.lri.fr/gallery/multiprecision.en.html.

Case study: algorithms for matrix multiplication. M. Clochard, L. Gondelman and M. Pereira worked on a case study about matrix multiplication. Two variants for the multiplication of matrices are proved: a naive version using three nested loops and Strassen's algorithm. To formally specify the two multiplication algorithms, they developed a new Why3 theory of matrices, and they applied a reflection methodology to conduct some of the proofs. A first version of this work was presented at the VSTTE Conference in 2016 [78]. An extended version that considers arbitrary rectangular matrices instead of square ones is published in the Journal of Automated Reasoning [12]. The development is available in Toccata's gallery http://toccata.lri.fr/gallery/verifythis_2016_matrix_multiplication.en.html.

Case studies: Strongly Connected Components in Directed Graphs As part of a larger set of case studies on algorithms on graphs http://pauillac.inria.fr/~levy/why3/, R. Chen and J.-J. Lévy work on formal verification of algorithms for computing strongly connected components of directed graphs.

The formal proofs are conducted using Why3. The formal proof of Tarjan's algorithm was presented at the French-speaking symposium JFLA 2017 [25] and then at the VSTTE 2017 international conference [17]

A Formally Proved, Complete Algorithm for Path Resolution with Symbolic Links In the context of file systems like those of Unix, path resolution is the operation that given a character string denoting an access path, determines the target object (a file, a directory, etc.) designated by this path. This operation is not trivial because of the presence of symbolic links. Indeed, the presence of such links may induce infinite loops in the resolution process. R. Chen, M. Clochard and C. Marché consider a path resolution algorithm that always terminates, detecting if it enters an infinite loop and reports a resolution failure in such a case. They propose a formal specification of path resolution and they formally prove that their algorithm terminates on any input, and is correct and complete with respect to this formal specification. [11]. The development is available in Toccata's gallery http://toccata.lri.fr/gallery/path_resolution.en.html.

7.4. Floating-Point and Numerical Programs

Computer Arithmetic and Formal Proofs: Verifying Floating-point Algorithms with the Coq System S. Boldo and G. Melquiond published a book that provides a comprehensive view of how to formally specify and verify tricky floating-point algorithms with the Coq proof assistant. It describes the Flocq formalization of floating-point arithmetic and some methods to automate theorem proofs. It then presents the specification and verification of various algorithms, from error-free transformations to a numerical scheme for a partial differential equation. The examples cover not only mathematical algorithms but also C programs as well as issues related to compilation [32].

Automating the Verification of Floating-Point Programs. The level of proof success and proof automation highly depends on the way the floating-point operations are interpreted in the logic supported by back-end provers. C. Fumex, C. Marché and Y. Moy addressed this challenge by combining multiple techniques to separately prove different parts of the desired properties. They use abstract interpretation to compute numerical bounds of expressions, and use multiple automated provers, relying on different strategies for representing floating-point computations. One of these strategies is based on the native support for floating-point arithmetic recently added in the SMT-LIB standard. The approach is implemented in the Why3 environment and its front-end SPARK 2014. It is validated experimentally on several examples originating from industrial use of SPARK 2014 [37], [21].

Round-off Error Analysis of Explicit One-Step Numerical Integration Methods. S. Boldo, A. Chapoutot, and F. Faissole provided bounds on the round-off errors of explicit one-step numerical integration methods, such as Runge-Kutta methods. They developed a fine-grained analysis that takes advantage of the linear stability of the scheme, a mathematical property that vouches the scheme is well-behaved [14].

Robustness of 2Sum and Fast2Sum. S. Boldo, S.Graillat, and J.-M. Muller worked on the 2Sum and Fast2Sum algorithms, that are important building blocks in numerical computing. They are used (implicitly or explicitly) in many compensated algorithms or for manipulating floating-point expansions. They showed that these algorithms are much more robust than it is usually believed: the returned result makes sense even when the rounding function is not round-to-nearest, and they are almost immune to overflow [10].

Formal Verification of a Floating-Point Expansion Renormalization Algorithm. Many numerical problems require a higher computing precision than the one offered by standard floating-point formats. A common way of extending the precision is to use floating-point expansions. S. Boldo, M. Joldes, J.-M. Muller, and V. Popescu proved one of the algorithms used as a basic brick when computing with floating-point expansions: renormalization that "compresses" an expansion [15].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. ProofInUse Joint Laboratory

Participants: Claude Marché [contact], Jean-Christophe Filliâtre, Andrei Paskevich.

ProofInUse is a joint project between the Toccata team and the SME AdaCore. It was selected and funded by the ANR programme “Laboratoires communs”, starting from April 2014, for 3 years http://www.spark-2014.org/proofinuse.

The SME AdaCore is a software publisher specializing in providing software development tools for critical systems. A previous successful collaboration between Toccata and AdaCore enabled *Why3* technology to be put into the heart of the AdaCore-developed SPARK technology.

The goal is now to promote and transfer the use of deduction-based verification tools to industry users, who develop critical software using the programming language Ada. The proof tools are aimed at replacing or complementing the existing test activities, whilst reducing costs.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. ELEFFAN

Participant: Sylvie Boldo [contact].

ELEFFAN is a Digicosme project funding the PhD of F. Faissole. S. Boldo is the principal investigator. It began in 2016 for three years. https://project.inria.fr/eleffan/

The ELEFFAN project aims at formally proving rounding error bounds of numerical schemes.

Partners: ENSTA Paristech (A. Chapoutot)

9.2. National Initiatives

9.2.1. ANR CoLiS

Participants: Claude Marché [contact], Andrei Paskevich.

The CoLiS research project is funded by the programme “Société de l’information et de la communication” of the ANR, for a period of 60 months, starting on October 1st, 2015. http://colis.irif.univ-paris-diderot.fr/

The project aims at developing formal analysis and verification techniques and tools for scripts. These scripts are written in the POSIX or bash shell language. Our objective is to produce, at the end of the project, formal methods and tools allowing to analyze, test, and validate scripts. For this, the project will develop techniques and tools based on deductive verification and tree transducers stemming from the domain of XML documents.

Partners: Université Paris-Diderot, IRIF laboratory (formerly PPS & LIAFA), coordinator; Inria Lille, team LINKS

9.2.2. ANR Vocal

Participants: Jean-Christophe Filliâtre [contact], Andrei Paskevich.

The Vocal research project is funded by the programme “Société de l’information et de la communication” of the ANR, for a period of 60 months, starting on October 1st, 2015. https://vocal.lri.fr/

The goal of the Vocal project is to develop the first formally verified library of efficient general-purpose data structures and algorithms. It targets the OCaml programming language, which allows for fairly efficient code and offers a simple programming model that eases reasoning about programs. The library will be readily available to implementers of safety-critical OCaml programs, such as Coq, Astrée, or Frama-C. It will provide the essential building blocks needed to significantly decrease the cost of developing safe software. The project intends to combine the strengths of three verification tools, namely Coq, Why3, and CFML. It will use Coq to obtain a common mathematical foundation for program specifications, as well as to verify purely functional components. It will use Why3 to verify a broad range of imperative programs with a high degree of proof automation. Finally, it will use CFML for formal reasoning about effectful higher-order functions and data structures making use of pointers and sharing.

Partners: team Gallium (Inria Paris-Rocquencourt), team DCS (Verimag), TrustInSoft, and OCamlPro.

9.2.3. ANR FastRelax

Participants: Sylvie Boldo [contact], Guillaume Melquiond.

This is a research project funded by the programme “Ingénierie Numérique & Sécurité” of the ANR. It is funded for a period of 48 months and it has started on October 1st, 2014. http://fastrelax.gforge.inria.fr/

Our aim is to develop computer-aided proofs of numerical values, with certified and reasonably tight error bounds, without sacrificing efficiency. Applications to zero-finding, numerical quadrature or global optimization can all benefit from using our results as building blocks. We expect our work to initiate a "fast and reliable" trend in the symbolic-numeric community. This will be achieved by developing interactions between our fields, designing and implementing prototype libraries and applying our results to concrete problems originating in optimal control theory.

Partners: team ARIC (Inria Grenoble Rhône-Alpes), team MARELLE (Inria Sophia Antipolis - Méditerranée), team SPECFUN (Inria Saclay - Île-de-France), Université Paris 6, and LAAS (Toulouse).

9.2.4. ANR Soprano

Participants: Sylvain Conchon [contact], Guillaume Melquiond.

The Soprano research project is funded by the programme “Sciences et technologies logicielles” of the ANR, for a period of 42 months, starting on October 1st, 2014. http://soprano-project.fr/

The SOPRANO project aims at preparing the next generation of verification-oriented solvers by gathering experts from academia and industry. We will design a new framework for the cooperation of solvers, focused on model generation and borrowing principles from SMT (current standard) and CP (well-known in optimization). Our main scientific and technical objectives are the following. The first objective is to design a new collaboration framework for solvers, centered around synthesis rather than satisfiability and allowing cooperation beyond that of Nelson-Oppen while still providing minimal interfaces with theoretical guarantees. The second objective is to design new decision procedures for industry-relevant and hard-to-solve theories. The third objective is to implement these results in a new open-source platform. The fourth objective is to ensure industrial-adequacy of the techniques and tools developed through periodical evaluations from the industrial partners.

Partners: team DIVERSE (Inria Rennes - Bretagne Atlantique), Adacore, CEA List, Université Paris-Sud, and OCamlPro.

9.2.5. FUI LCHIP

Participant: Sylvain Conchon [contact].

LCHIP (Low Cost High Integrity Platform) is aimed at easing the development of safety critical applications (up to SIL4) by providing: (i) a complete IDE able to automatically generate and prove bounded complexity software (ii) a low cost, safe execution platform. The full support of DSLs and third party code generators will enable a seamless deployment into existing development cycles. LCHIP gathers scientific results obtained during the last 20 years in formal methods, proof, refinement, code generation, etc. as well as a unique return of experience on safety critical systems design. http://www.clearsy.com/en/2016/10/4260/

Partners: 2 technology providers (ClearSy, OcamlPro), in charge of building the architecture of the platform; 3 labs (IFSTTAR, LIP6, LRI), to improve LCHIP IDE features; 2 large companies (SNCF, RATP), representing public ordering parties, to check compliance with standard and industrial railway use-case.

The project lead by ClearSy has started in April 2016 and lasts 3 years. It is funded by BpiFrance as well as French regions.

9.2.6. ANR PARDI

Participant: Sylvain Conchon [contact].

Verification of PARameterized DIstributed systems. A parameterized system specification is a specification for a whole class of systems, parameterized by the number of entities and the properties of the interaction, such as the communication model (synchronous/asynchronous, order of delivery of message, application ordering) or the fault model (crash failure, message loss). To assist and automate verification without parameter instantiation, PARDI uses two complementary approaches. First, a fully automatic model checker modulo theories is considered. Then, to go beyond the intrinsic limits of parameterized model checking, the project advocates a collaborative approach between proof assistant and model checker. http://pardi.enseeiht.fr/

The proof lead by Toulouse INP/IRIT started in 2016 and lasts for 4 years. Partners: Université Pierre et Marie Curie (LIP6), Université Paris-Sud (LRI), Inria Nancy (team VERIDIS)

9.3. European Initiatives

9.3.1. Collaborations in European Programs, Except FP7 & H2020

Program: COST (European Cooperation in Science and Technology).

Project acronym: EUTypes https://eutypes.cs.ru.nl/

Project title: The European research network on types for programming and verification

Duration: 2015-2019

Coordinator: Herman Geuvers, Radboud University Nijmegen, The Netherlands

Other partners: 36 members countries, see http://www.cost.eu/COST_Actions/ca/CA15123?parties

Abstract: Types are pervasive in programming and information technology. A type defines a formal interface between software components, allowing the automatic verification of their connections, and greatly enhancing the robustness and reliability of computations and communications. In rich dependent type theories, the full functional specification of a program can be expressed as a type. Type systems have rapidly evolved over the past years, becoming more sophisticated, capturing new aspects of the behaviour of programs and the dynamics of their execution.

This COST Action will give a strong impetus to research on type theory and its many applications in computer science, by promoting (1) the synergy between theoretical computer scientists, logicians and mathematicians to develop new foundations for type theory, for example as based on the recent development of "homotopy type theory", (2) the joint development of type theoretic tools as proof assistants and integrated programming environments, (3) the study of dependent types for programming and its deployment in software development, (4) the study of dependent types for verification and its deployment in software analysis and verification. The action will also tie together these different areas and promote cross-fertilisation.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

Ran Chen is a PhD student from Institute of Software (Chinese Academy of Sciences, Beijing, China) visiting the team for 10 months under the supervision of C. Marché and J.-J. Lévy (PiR2 team, Inria Paris). She worked on the formal verification of graphs algorithms [25], [17], and also in the context of the CoLiS project on verification of some aspects of the Unix file system and shell scripts [74] [11]

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

S. Boldo, vice-president of the 28th “Journées Francophones des Langages Applicatifs” (JFLA 2017)

S. Boldo, president of the 29th “Journées Francophones des Langages Applicatifs” (JFLA 2018)

J.-C. Filliâtre, scientific chair and co-organizer of EJCP (École Jeunes Chercheurs en Programmation du GDR GPL) at Toulouse on June 26–30, 2017. http://ejcp2017.enseeiht.fr/

10.1.1.2. Member of the Organizing Committees

S. Conchon, local chair for the 44th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL 2017), held in Paris, France in January 2017. http://conf.researchr.org/home/POPL-2017).

C. Marché, co-organizer of the first joint Frama-C/SPARK day (May, Paris, http://frama-c.com/FCSD17.html), in the context of the Open Source Innovation Spring (http://www.open-source-innovation-spring.org/).

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

A. Paskevich, program chair of the 9th Working Conference on Verified Software: Theories, Tools, and Experiments (VSTTE 2017), in collaboration with Thomas Wies (NYU) [35].

S. Boldo, program chair of the 10th International Workshop on Numerical Software Verification (NSV 2017) in collaboration with Alessandro Abate (Oxford) [33].

S. Boldo, program vice-chair of the 28th “Journées Francophones des Langages Applicatifs” (JFLA 2017) [34].

S. Boldo, program chair of the 29th “Journées Francophones des Langages Applicatifs” (JFLA 2018).

10.1.2.2. Member of the Conference Program Committees

S. Boldo, PC of the 24th IEEE Symposium on Computer Arithmetic (ARITH 2017)

S. Boldo, PC of the 25th IEEE Symposium on Computer Arithmetic (ARITH 2018)

S. Boldo, PC of the 6th ACM SIGPLAN Conference on Certified Programs and Proofs (CPP 2017)

S. Boldo, PC of the 7th ACM SIGPLAN Conference on Certified Programs and Proofs (CPP 2018)

S. Boldo, PC of the 8th International Conference on Interactive Theorem Proving (ITP 2017)

S. Boldo, PC of the Tenth NASA Formal Methods Symposium (NFM 2018)

G. Melquiond, PC of the 3rd International Workshop on Coq for Programming Languages (CoqPL 2017).

G. Melquiond, PC of the 1st ACM SIGPLAN Workshop on Machine Learning and Programming Languages (MAPL 2017).

G. Melquiond, PC of the 10th International Workshop on Numerical Software Verification (NSV 2017).

10.1.2.3. Reviewer

The members of the Toccata team have reviewed papers for numerous international conferences.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

G. Melquiond, member of the editorial board of *Reliable Computing*.

S. Boldo, member of the editorial board of Binaire http://binaire.blog.lemonde.fr, the blog of the French Computer Science Society.

10.1.3.2. Reviewer - Reviewing Activities

The members of the Toccata team have reviewed numerous papers for numerous international journals.

10.1.4. Invited Talks

S. Boldo gave a talk at EDF in Palaiseau on April 20th

S. Boldo gave a talk at the ModeliScale IPL in Paris on July 4th

S. Boldo gave a talk to teachers in Luminy on May 4th

S. Boldo gave a talk at the université of Saint-Denis de la Réunion on December 8th

10.1.5. Leadership within the Scientific Community

S. Boldo, elected chair of the ARITH working group of the GDR-IM (a CNRS subgroup of computer science) with J. Detrey (Inria Nancy).

10.1.6. Scientific Expertise

C. Marché, member of the scientific commission of Inria-Saclay, in charge of selecting candidates for PhD grants, Post-doc grants, temporary leaves from universities ("délégations").

C. Marché, member of the "Bureau du Comité des Projets" of Inria-Saclay, in charge of examining proposals for creation of new Inria project-teams.

S. Boldo, member of the program committee for selecting postdocs of the maths/computer science program of the Labex mathématique Hadamard.

S. Boldo, member of a hiring committee for an associate professor position in computer science at Université Joseph Fourier, Grenoble, France.

S. Boldo, member of the 2017 committee for the Gilles Kahn PhD award of the French Computer Science Society.

10.1.7. Research Administration

G. Melquiond, member of the committee for the monitoring of PhD students (*"commission de suivi doctoral"*).

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master Parisien de Recherche en Informatique (MPRI) https://wikimpri.dptinfo.ens-cachan.fr/doku.php: "Proofs of Programs" http://www.lri.fr/~marche/MPRI-2-36-1/ (M2), C. Marché (12h), A. Charguéraud (12h), Université Paris-Diderot, France.

Master: Fondements de l'informatique et ingénierie du logiciel (FIIL) https://www.lri.fr/~conchon/parcours_fiil/: "Software Model Checking" (M2), S. Conchon (9h), "Programmation C++11 avancée" (M2), G. Melquiond (16h), "Vérification déductive de programmes" (M2), A. Paskevich (10.5h), Université Paris-Sud, France.

DUT (Diplôme Universitaire de Technologie): M1101 "Introduction aux systèmes informatiques", A. Paskevich (36h), M3101 "Principes des systèmes d'exploitation", A. Paskevich (58.5h), IUT d'Orsay, Université Paris-Sud, France.

Licence: “Langages de programmation et compilation” (L3), J.-C. Filliâtre (26h), École Normale Supérieure, France.

Licence: “INF411: Les bases de l’algorithmique et de la programmation” (L3), J.-C. Filliâtre (16h), École Polytechnique, France.

Master: “INF564: Compilation” (M1), J.-C. Filliâtre (18h), École Polytechnique, France.

Licence: “Programmation fonctionnelle avancée” (L3), S. Conchon (45h), Université Paris-Sud, France.

Licence: “Introduction à la programmation fonctionnelle” (L2), S. Conchon (25h), Université Paris-Sud, France.

10.2.2. Internships

R. Rieu-Helft (ENS, Paris) was a pre-PhD student doing an internship under supervision of C. Marché and G. Melquiond. He worked on the design and the formal verification of a library for unbounded integer arithmetic [23]. He implemented in Why3 a mechanism for extracting code to the C language, in order to obtain a certified code that runs very efficiently [30].

D. Gallois-Wong was a Master-2 intern for 4 months under the supervision of S. Boldo. She began the formalization in Coq of numerical filters.

V. Tourneur was a Master-1 intern for 4 months under the supervision of S. Boldo. He developed and proved a new algorithm for computing the average of two floating-point numbers when the radix is 10.

10.2.3. Supervision

PhD in progress: M. Clochard, “Méthodes et outils pour la spécification et la preuve de propriétés difficiles de programmes séquentiels”, since Oct. 2013, supervised by C. Marché and A. Paskevich.

PhD in progress: D. Declerck, “Vérification par des techniques de test et model checking de programmes C11”, since Sep. 2014, supervised by F. Zaïdi (LRI) and S. Conchon.

PhD in progress: M. Roux, “Model Checking de systèmes paramétrés et temporisés”, since Sep. 2015, supervised by Sylvain Conchon.

PhD in progress: M. Pereira, “A Verified Graph Library. Tools and techniques for the verification of modular higher-order programs, with extraction”, since May 2015, supervised by J.-C. Filliâtre.

PhD in progress: A. Coquereau, “[ErgoFast] Amélioration de performances pour le solveur SMT Alt-Ergo : conception d’outils d’analyse, optimisations et structures de données efficaces pour OCaml”, since Sep. 2015, supervised by S. Conchon, F. Le Fessant et M. Mauny.

PhD in progress: F. Faissole, “Stabilité(s): liens entre l’arithmétique flottante et l’analyse numérique”, since Oct. 2016, supervised by S. Boldo and A. Chapoutot.

PhD in progress: R. Rieu-Helft, “Développement et vérification de bibliothèques d’arithmétique entière en précision arbitraire”, since Oct. 2017, supervised by G. Melquiond and P. Cuoq (TrustIn-Soft).

PhD in progress: D. Gallois-Wong, “Vérification formelle et filtres numériques”, since Oct. 2017, supervised by S. Boldo and T. Hilaire.

10.2.4. Juries

C. Marché: reviewer of the habilitation thesis of R. Bubel, “Deductive Verification: From Theory to Practice”, Technische Universität Darmstadt, Germany, November 2017.

S. Boldo: reviewer and member of the PhD defense of A. Plet, École Normale Supérieure de Lyon, Lyon, France, July 2017.

S. Boldo: reviewer and member of the PhD defense of F. Maurica, Université de la Réunion, Saint-Denis, France, December 2017.

S. Boldo: president of the PhD defense of T. Sibut-Pinote, Université Paris-Saclay, Palaiseau, France, December 2017.

10.3. Popularization

S. Boldo, scientific head for Saclay for the MECSI group for networking about computer science popularization inside Inria.

S. Boldo gave a talk at the Inria Saclay about how to popularize programming.

During the "Fête de la science" on October 13th, S. Boldo demonstrated unplugged computer science to teenagers and F. Faissole run a stand about an introduction to programming with robots. S. Boldo also did this activity to kids from 7 to 17 at the Massy opera on November, 17th.

S. Boldo gave a talk during at a *Girls can code* weekend on August 23rd in Paris.

S. Boldo went to the Arpajon high-school for presenting Women in Science on December 19th.

S. Boldo gave a popularization talk to the administrative staff of Inria at Rocquencourt for the Inria birthday on November 16th.

11. Bibliography

Major publications by the team in recent years

[1] F. BOBOT, S. CONCHON, É. CONTEJEAN, M. IGUERNELALA, A. MAHBOUBI, A. MEBSOUT, G. MELQUIOND.*A Simplex-Based Extension of Fourier-Motzkin for Solving Linear Integer Arithmetic*, in "IJCAR 2012: Proceedings of the 6th International Joint Conference on Automated Reasoning", Manchester, UK, B. GRAMLICH, D. MILLER, U. SATTLER (editors), Lecture Notes in Computer Science, Springer, June 2012, vol. 7364, p. 67–81, http://hal.inria.fr/hal-00687640.

[2] F. BOBOT, J.-C. FILLIÂTRE, C. MARCHÉ, A. PASKEVICH.*Let's Verify This with Why3*, in "International Journal on Software Tools for Technology Transfer (STTT)", 2015, vol. 17, n^{o} 6, p. 709–727, http://hal.inria.fr/hal-00967132/en.

[3] S. BOLDO, F. CLÉMENT, J.-C. FILLIÂTRE, M. MAYERO, G. MELQUIOND, P. WEIS.*Wave Equation Numerical Resolution: a Comprehensive Mechanized Proof of a C Program*, in "Journal of Automated Reasoning", April 2013, vol. 50, n^{o} 4, p. 423–456, http://hal.inria.fr/hal-00649240/en/.

[4] S. BOLDO, G. MELQUIOND.*Flocq: A Unified Library for Proving Floating-point Algorithms in Coq*, in "Proceedings of the 20th IEEE Symposium on Computer Arithmetic", Tübingen, Germany, E. ANTELO, D. HOUGH, P. IENNE (editors), 2011, p. 243–252, http://hal.archives-ouvertes.fr/inria-00534854/.

[5] S. CONCHON, A. GOEL, S. KRSTIĆ, A. MEBSOUT, F. ZAÏDI.*Cubicle: A Parallel SMT-based Model Checker for Parameterized Systems*, in "CAV 2012: Proceedings of the 24th International Conference on Computer Aided Verification", Berkeley, California, USA, M. PARTHASARATHY, S. A. SESHIA (editors), Lecture Notes in Computer Science, Springer, July 2012, vol. 7358, http://hal.archives-ouvertes.fr/hal-00799272.

[6] J.-C. FILLIÂTRE, L. GONDELMAN, A. PASKEVICH.*The Spirit of Ghost Code*, in "Formal Methods in System Design", 2016, vol. 48, n^{o} 3, p. 152–174, https://hal.archives-ouvertes.fr/hal-01396864v1.

[7] C. FUMEX, C. DROSS, J. GERLACH, C. MARCHÉ.*Specification and Proof of High-Level Functional Properties of Bit-Level Programs*, in "8th NASA Formal Methods Symposium", Minneapolis, MN, USA, S. RAYADURGAM, O. TKACHUK (editors), Lecture Notes in Computer Science, Springer, June 2016, vol. 9690, p. 291–306, https://hal.inria.fr/hal-01314876.

[8] C. MARCHÉ.*Verification of the Functional Behavior of a Floating-Point Program: an Industrial Case Study*, in "Science of Computer Programming", March 2014, vol. 96, n° 3, p. 279–296, http://hal.inria.fr/hal-00967124/en.

[9] É. MARTIN-DOREL, G. MELQUIOND.*Proving Tight Bounds on Univariate Expressions with Elementary Functions in Coq*, in "Journal of Automated Reasoning", 2016, https://hal.inria.fr/hal-01086460.

Publications of the year

Articles in International Peer-Reviewed Journal

[10] S. BOLDO, S. GRAILLAT, J.-M. MULLER.*On the robustness of the 2Sum and Fast2Sum algorithms*, in "ACM Transactions on Mathematical Software", July 2017, vol. 44, n° 1, https://hal-ens-lyon.archives-ouvertes.fr/ensl-01310023.

[11] R. CHEN, M. CLOCHARD, C. MARCHÉ.*A Formally Proved, Complete Algorithm for Path Resolution with Symbolic Links*, in "Journal of Formalized Reasoning", November 2017, vol. 10, n° 1, https://hal.inria.fr/hal-01652148.

[12] M. CLOCHARD, L. GONDELMAN, M. PEREIRA.*The Matrix Reproved: Verification Pearl*, in "Journal of Automated Reasoning", October 2017, p. 1–19 [*DOI :* 10.1007/S10817-017-9436-2], https://hal.inria.fr/hal-01617437.

International Conferences with Proceedings

[13] S. BOLDO, F. CLÉMENT, F. FAISSOLE, V. MARTIN, M. MAYERO.*A Coq Formal Proof of the Lax–Milgram theorem*, in "6th ACM SIGPLAN Conference on Certified Programs and Proofs", Paris, France, January 2017 [*DOI :* 10.1145/3018610.3018625], https://hal.inria.fr/hal-01391578.

[14] S. BOLDO, F. FAISSOLE, A. CHAPOUTOT.*Round-off Error Analysis of Explicit One-Step Numerical Integration Methods*, in "24th IEEE Symposium on Computer Arithmetic", London, United Kingdom, July 2017 [*DOI :* 10.1109/ARITH.2017.22], https://hal.archives-ouvertes.fr/hal-01581794.

[15] S. BOLDO, M. JOLDES, J.-M. MULLER, V. POPESCU.*Formal Verification of a Floating-Point Expansion Renormalization Algorithm*, in "8th International Conference on Interactive Theorem Proving (ITP'2017)", Brasilia, Brazil, Lecture Notes in Computer Science, September 2017, vol. 10499, https://hal.archives-ouvertes.fr/hal-01512417.

[16] A. CHARGUÉRAUD, F. POTTIER.*Temporary Read-Only Permissions for Separation Logic*, in "Proceedings of the 26th European Symposium on Programming (ESOP 2017)", Uppsala, Sweden, April 2017, https://hal.inria.fr/hal-01408657.

[17] R. CHEN, J.-J. LÉVY.*A Semi-automatic Proof of Strong connectivity*, in "9th Working Conference on Verified Software: Theories, Tools and Experiments (VSTTE)", Heidelberg, Germany, July 2017, https://hal.inria.fr/hal-01632947.

[18] S. CONCHON, M. IGUERNELALA, K. JI, G. MELQUIOND, C. FUMEX.*A Three-tier Strategy for Reasoning about Floating-Point Numbers in SMT*, in "29th International Conference on Computer Aided Verification", Heidelberg, Germany, V. KUNCAK, R. MAJUMDAR (editors), Lecture Notes in Computer Science,

Springer, July 2017, vol. 10427, p. 419-435 [*DOI :* 10.1007/978-3-319-63390-9_22], https://hal.inria.fr/hal-01522770.

[19] F. FAISSOLE.*Formalization and closedness of finite dimensional subspaces*, in "SYNASC 2017 - 19th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing", Timioara, Romania, September 2017, p. 1-7, https://hal.inria.fr/hal-01630411.

[20] J.-C. FILLIÂTRE, M. PEREIRA, S. MELO DE SOUSA.*Vérification de programmes OCaml fortement impératifs avec Why3*, in "JFLA 2018 - Journées Francophones des Langages Applicatifs", Banyuls-sur-Mer, France, January 2018, p. 1-14, https://hal.inria.fr/hal-01649989.

[21] C. FUMEX, C. MARCHÉ, Y. MOY.*Automating the Verification of Floating-Point Programs*, in "9th Working Conference on Verified Software: Theories, Tools and Experiments", Heidelberg, Germany, A. PASKEVICH, T. WIES (editors), Lecture Notes in Computer Science, Springer, July 2017, vol. 10712, https://hal.inria.fr/hal-01534533.

[22] N. JEANNEROD, C. MARCHÉ, R. TREINEN.*A Formally Verified Interpreter for a Shell-like Programming Language*, in "VSTTE 2017 - 9th Working Conference on Verified Software: Theories, Tools, and Experiments", Heidelberg, Germany, Lecture Notes in Computer Science, July 2017, vol. 10712, https://hal.archives-ouvertes.fr/hal-01534747.

[23] R. RIEU-HELFT, C. MARCHÉ, G. MELQUIOND.*How to Get an Efficient yet Verified Arbitrary-Precision Integer Library*, in "9th Working Conference on Verified Software: Theories, Tools, and Experiments", Heidelberg, Germany, A. PASKEVICH, T. WIES (editors), Lecture Notes in Computer Science, July 2017, vol. 10712, p. 84-101 [*DOI :* 10.1007/978-3-319-72308-2_6], https://hal.inria.fr/hal-01519732.

National Conferences with Proceeding

[24] S. BOLDO, F. CLÉMENT, F. FAISSOLE, V. MARTIN, M. MAYERO.*Preuve formelle du théorème de Lax–Milgram*, in "16èmes journées Approches Formelles dans l'Assistance au Développement de Logiciels", Montpellier, France, June 2017, https://hal.archives-ouvertes.fr/hal-01581807.

[25] R. CHEN, J.-J. LÉVY.*Une preuve formelle de l'algorithme de Tarjan-1972 pour trouver les composantes fortement connexes dans un graphe*, in "JFLA 2017 - Vingt-huitièmes Journées Francophones des Langages Applicatifs", Gourette, France, Vingt-huitièmes Journées Francophones des Langages Applicatifs, January 2017, https://hal.inria.fr/hal-01422215.

[26] M. CLOCHARD.*Preuves taillées en biseau*, in "vingt-huitièmes Journées Francophones des Langages Applicatifs (JFLA)", Gourette, France, January 2017, https://hal.inria.fr/hal-01404935.

[27] F. FAISSOLE.*Définir le fini : deux formalisations d'espaces de dimension finie*, in "JLFA 2017 - Journées Francophones des Langages Applicatifs", Banyuls-sur-mer, France, 29èmes Journées Francophones des Langages Applicatifs, January 2018, p. 1-6, https://hal.inria.fr/hal-01654457.

[28] F. FAISSOLE, B. SPITTERS.*Preuves constructives de programmes probabilistes*, in "JFLA 2018 - Journées Francophones des Langages Applicatifs", Banyuls-sur-Mer, France, 29èmes Journées Francophones des Langages Applicatifs, January 2018, https://hal.inria.fr/hal-01654459.

[29] M. PEREIRA.*Défonctionnaliser pour prouver*, in "JFLA 2017 - Vingt-huitième Journées Francophones des Langages Applicatifs", Gourette, France, January 2017, https://hal.inria.fr/hal-01378068.

[30] R. RIEU-HELFT.*Un mécanisme d'extraction vers C pour Why3*, in "29èmes Journées Francophones des Langages Applicatifs", Banyuls-sur-Mer, France, January 2018, https://hal.inria.fr/hal-01653153.

Conferences without Proceedings

[31] F. FAISSOLE, B. SPITTERS.*Synthetic topology in HoTT for probabilistic programming*, in "The Third International Workshop on Coq for Programming Languages (CoqPL 2017)", Paris, France, January 2017, https://hal.inria.fr/hal-01405762.

Scientific Books (or Scientific Book chapters)

[32] S. BOLDO, G. MELQUIOND.*Computer Arithmetic and Formal Proofs: Verifying Floating-point Algorithms with the Coq System*, ISTE Press - Elsevier, December 2017, 326, https://hal.inria.fr/hal-01632617.

Books or Proceedings Editing

[33] A. ABATE, S. BOLDO (editors). *10th International Workshop on Numerical Software Verification*, Springer, July 2017, https://hal.inria.fr/hal-01662076.

[34] S. BOLDO, J. SIGNOLES (editors). *Vingt-huitièmes Journées Francophones des Langages Applicatifs*, Published by the authors, January 2017, https://hal.inria.fr/hal-01662072.

[35] A. PASKEVICH, T. WIES (editors). *Verified Software: Theories, Tools, and Experiments, Revised Selected Papers Presented at the 9th International Conference VSTTE*, springer, Heidelberg, Germany, December 2017, vol. Lecture Notes in Computer Science, n° 10712 [*DOI :* 10.1007/978-3-319-72308-2], https://hal.inria.fr/hal-01670145.

Research Reports

[36] I. DAMI, C. MARCHÉ.*The CoLiS language: syntax, semantics and associated tools*, Inria Saclay Ile de France, October 2017, n° RT-0491, p. 1-22, https://hal.inria.fr/hal-01614488.

[37] C. FUMEX, C. MARCHÉ, Y. MOY.*Automated Verification of Floating-Point Computations in Ada Programs*, Inria Saclay Ile de France, April 2017, n° RR-9060, 53, https://hal.inria.fr/hal-01511183.

Other Publications

[38] A. CHARGUÉRAUD, J.-C. FILLIÂTRE, M. PEREIRA, F. POTTIER.*VOCAL – A Verified OCAml Library*, September 2017, ML Family Workshop 2017, https://hal.inria.fr/hal-01561094.

[39] F. FAISSOLE, B. SPITTERS.*Synthetic topology in homotopy type theory for probabilistic programming*, January 2017, 3, PPS 2017 - Workshop on probabilistic programming semantics , Poster, https://hal.inria.fr/hal-01485397.

[40] A. MAHBOUBI, G. MELQUIOND, T. SIBUT-PINOTE.*Formally Verified Approximations of Definite Integrals*, February 2017, working paper or preprint, https://hal.inria.fr/hal-01630143.

References in notes

[41] B. BECKERT, R. HÄHNLE, P. H. SCHMITT (editors). *Verification of Object-Oriented Software: The KeY Approach*, Lecture Notes in Computer Science, Springer, 2007, vol. 4334.

[42] U. A. ACAR, A. CHARGUÉRAUD, M. RAINEY.*Theory and Practice of Chunked Sequences*, in "European Symposium on Algorithms", Wroclaw, Poland, A. SCHULZ, D. WAGNER (editors), Springer, September 2014, vol. Lecture Notes in Computer Science, n^{o} 8737, p. 25–36, https://hal.inria.fr/hal-01087245.

[43] U. A. ACAR, A. CHARGUÉRAUD, M. RAINEY.*Oracle-Guided Scheduling for Controlling Granularity in Implicitly Parallel Languages*, in "Journal of Functional Programming", November 2016, vol. 26, https://hal.inria.fr/hal-01409069.

[44] U. A. ACAR, A. CHARGUÉRAUD, M. RAINEY, F. SIECZKOWSKI.*Dag-calculus: a calculus for parallel computation*, in "Proceedings of the 21st ACM SIGPLAN International Conference on Functional Programming (ICFP)", Nara, Japan, September 2016, p. 18–32, https://hal.inria.fr/hal-01409022.

[45] J. B. ALMEIDA, M. BARBOSA, J.-C. FILLIÂTRE, J. S. PINTO, B. VIEIRA.*CAOVerif: An Open-Source Deductive Verification Platform for Cryptographic Software Implementations*, in "Science of Computer Programming", October 2012.

[46] A. AYAD, C. MARCHÉ.*Multi-Prover Verification of Floating-Point Programs*, in "Fifth International Joint Conference on Automated Reasoning", Edinburgh, Scotland, J. GIESL, R. HÄHNLE (editors), Lecture Notes in Artificial Intelligence, Springer, July 2010, vol. 6173, p. 127–141, http://hal.inria.fr/inria-00534333.

[47] D. BAELDE, P. COURTIEU, D. GROSS-AMBLARD, C. PAULIN-MOHRING.*Towards Provably Robust Watermarking*, in "ITP 2012", Lecture Notes in Computer Science, August 2012, vol. 7406, http://hal.inria.fr/hal-00682398.

[48] C. BARRETT, C. TINELLI.*CVC3*, in "19th International Conference on Computer Aided Verification", Berlin, Germany, W. DAMM, H. HERMANNS (editors), Lecture Notes in Computer Science, Springer, July 2007, vol. 4590, p. 298–302.

[49] P. BAUDIN, J.-C. FILLIÂTRE, C. MARCHÉ, B. MONATE, Y. MOY, V. PREVOSTO.*ACSL: ANSI/ISO C Specification Language, version 1.4*, 2009.

[50] P. BEHM, P. BENOIT, A. FAIVRE, J.-M. MEYNADIER.*METEOR : A successful application of B in a large project*, in "Proceedings of FM'99: World Congress on Formal Methods", J. M. WING, J. WOODCOCK, J. DAVIES (editors), Lecture Notes in Computer Science (Springer-Verlag), Springer Verlag, September 1999, p. 369–387.

[51] J. C. BLANCHETTE, A. PASKEVICH.*TFF1: The TPTP typed first-order form with rank-1 polymorphism*, in "24th International Conference on Automated Deduction (CADE-24)", Lake Placid, USA, Lecture Notes in Artificial Intelligence, Springer, June 2013, vol. 7898, http://hal.inria.fr/hal-00825086.

[52] F. BOBOT, S. CONCHON, É. CONTEJEAN, M. IGUERNELALA, S. LESCUYER, A. MEBSOUT.*The Alt-Ergo Automated Theorem Prover*, 2008.

[53] F. BOBOT, J.-C. FILLIÂTRE.*Separation Predicates: a Taste of Separation Logic in First-Order Logic*, in "14th International Conference on Formal Ingineering Methods (ICFEM)", Kyoto, Japan, Lecture Notes in Computer Science, Springer, November 2012, vol. 7635, http://hal.inria.fr/hal-00825088.

[54] F. BOBOT, J.-C. FILLIÂTRE, C. MARCHÉ, G. MELQUIOND, A. PASKEVICH.*Preserving User Proofs Across Specification Changes*, in "Verified Software: Theories, Tools, Experiments (5th International Conference VSTTE)", Atherton, USA, E. COHEN, A. RYBALCHENKO (editors), Lecture Notes in Computer Science, Springer, May 2013, vol. 8164, p. 191–201, http://hal.inria.fr/hal-00875395.

[55] F. BOBOT, J.-C. FILLIÂTRE, C. MARCHÉ, G. MELQUIOND, A. PASKEVICH.*The Why3 platform, version 0.81*, version 0.81, LRI, CNRS & Univ. Paris-Sud & Inria Saclay, March 2013, http://hal.inria.fr/hal-00822856/.

[56] M. BODIN, A. CHARGUÉRAUD, D. FILARETTI, P. GARDNER, S. MAFFEIS, D. NAUDZIUNIENE, A. SCHMITT, G. SMITH.*A Trusted Mechanised JavaScript Specification*, in "Proceedings of the 41st ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages", San Diego, USA, ACM Press, January 2014, http://hal.inria.fr/hal-00910135.

[57] S. BOLDO.*How to Compute the Area of a Triangle: a Formal Revisit*, in "Proceedings of the 21th IEEE Symposium on Computer Arithmetic", Austin, Texas, USA, 2013, http://hal.inria.fr/hal-00790071.

[58] S. BOLDO.*Deductive Formal Verification: How To Make Your Floating-Point Programs Behave*, Université Paris-Sud, October 2014, Thèse d'habilitation, https://hal.inria.fr/tel-01089643.

[59] S. BOLDO.*Formal verification of tricky numerical computations*, in "16th GAMM-IMACS International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics", Würzburg, Germany, September 2014, https://hal.inria.fr/hal-01088692.

[60] S. BOLDO, F. CLÉMENT, J.-C. FILLIÂTRE, M. MAYERO, G. MELQUIOND, P. WEIS.*Formal Proof of a Wave Equation Resolution Scheme: the Method Error*, in "Proceedings of the First Interactive Theorem Proving Conference", Edinburgh, Scotland, M. KAUFMANN, L. C. PAULSON (editors), LNCS, Springer, July 2010, vol. 6172, p. 147–162, http://hal.inria.fr/inria-00450789/.

[61] S. BOLDO, F. CLÉMENT, J.-C. FILLIÂTRE, M. MAYERO, G. MELQUIOND, P. WEIS.*Trusting Computations: a Mechanized Proof from Partial Differential Equations to Actual Program*, in "Computers and Mathematics with Applications", 2014, vol. 68, n^{o} 3, p. 325–352, http://hal.inria.fr/hal-00769201.

[62] S. BOLDO, J.-C. FILLIÂTRE, G. MELQUIOND.*Combining Coq and Gappa for Certifying Floating-Point Programs*, in "16th Symposium on the Integration of Symbolic Computation and Mechanised Reasoning", Grand Bend, Canada, Lecture Notes in Artificial Intelligence, Springer, July 2009, vol. 5625, p. 59–74.

[63] S. BOLDO, J.-H. JOURDAN, X. LEROY, G. MELQUIOND.*A Formally-Verified C Compiler Supporting Floating-Point Arithmetic*, in "Proceedings of the 21th IEEE Symposium on Computer Arithmetic", Austin, Texas, USA, 2013, http://hal.inria.fr/hal-00743090.

[64] S. BOLDO, J.-H. JOURDAN, X. LEROY, G. MELQUIOND.*Verified Compilation of Floating-Point Computations*, in "Journal of Automated Reasoning", February 2015, vol. 54, n^{o} 2, p. 135-163, https://hal.inria.fr/hal-00862689.

[65] S. BOLDO, C. LELAY, G. MELQUIOND.*Improving Real Analysis in Coq: a User-Friendly Approach to Integrals and Derivatives*, in "Proceedings of the Second International Conference on Certified Programs and Proofs", Kyoto, Japan, C. HAWBLITZEL, D. MILLER (editors), Lecture Notes in Computer Science, December 2012, vol. 7679, p. 289–304, http://hal.inria.fr/hal-00712938.

[66] S. BOLDO, C. LELAY, G. MELQUIOND.*Coquelicot: A User-Friendly Library of Real Analysis for Coq*, in "Mathematics in Computer Science", June 2015, vol. 9, n^o 1, p. 41-62, http://hal.inria.fr/hal-00860648.

[67] S. BOLDO, C. LELAY, G. MELQUIOND.*Formalization of Real Analysis: A Survey of Proof Assistants and Libraries*, in "Mathematical Structures in Computer Science", 2016, http://hal.inria.fr/hal-00806920.

[68] S. BOLDO, C. MARCHÉ.*Formal verification of numerical programs: from C annotated programs to mechanical proofs*, in "Mathematics in Computer Science", 2011, vol. 5, p. 377–393, http://hal.inria.fr/hal-00777605.

[69] S. BOLDO, T. M. T. NGUYEN.*Proofs of numerical programs when the compiler optimizes*, in "Innovations in Systems and Software Engineering", 2011, vol. 7, p. 151–160, http://hal.inria.fr/hal-00777639.

[70] T. BORMER, M. BROCKSCHMIDT, D. DISTEFANO, G. ERNST, J.-C. FILLIÂTRE, R. GRIGORE, M. HUISMAN, V. KLEBANOV, C. MARCHÉ, R. MONAHAN, W. MOSTOWSKI, N. POLIKARPOVA, C. SCHEBEN, G. SCHELLHORN, B. TOFAN, J. TSCHANNEN, M. ULBRICH.*The COST IC0701 Verification Competition 2011*, in "Formal Verification of Object-Oriented Software, Revised Selected Papers Presented at the International Conference, FoVeOOS 2011", B. BECKERT, F. DAMIANI, D. GUROV (editors), Lecture Notes in Computer Science, Springer, 2012, vol. 7421, http://hal.inria.fr/hal-00789525.

[71] L. BURDY, Y. CHEON, D. R. COK, M. D. ERNST, J. R. KINIRY, G. T. LEAVENS, K. R. M. LEINO, E. POLL.*An overview of JML tools and applications*, in "International Journal on Software Tools for Technology Transfer (STTT)", June 2005, vol. 7, n^o 3, p. 212–232.

[72] R. CHAPMAN, F. SCHANDA.*Are We There Yet? 20 Years of Industrial Theorem Proving with SPARK*, in "Interactive Theorem Proving - 5th International Conference, ITP 2014, Held as Part of the Vienna Summer of Logic, VSL 2014, Vienna, Austria, July 14-17, 2014. Proceedings", G. KLEIN, R. GAMBOA (editors), Lecture Notes in Computer Science, Springer, 2014, vol. 8558, p. 17–26.

[73] A. CHARGUÉRAUD, F. POTTIER.*Verifying the Correctness and Amortized Complexity of a Union-Find Implementation in Separation Logic with Time Credits*, in "Journal of Automated Reasoning", September 2017.

[74] R. CHEN, M. CLOCHARD, C. MARCHÉ.*A Formal Proof of a Unix Path Resolution Algorithm*, Inria, December 2016, n^o RR-8987, https://hal.inria.fr/hal-01406848.

[75] M. CLOCHARD.*Automatically verified implementation of data structures based on AVL trees*, in "6th Working Conference on Verified Software: Theories, Tools and Experiments (VSTTE)", Vienna, Austria, D. GIANNAKOPOULOU, D. KROENING (editors), Lecture Notes in Computer Science, Springer, July 2014, vol. 8471, p. 167–180, http://hal.inria.fr/hal-01067217.

[76] M. CLOCHARD, J.-C. FILLIÂTRE, C. MARCHÉ, A. PASKEVICH.*Formalizing Semantics with an Automatic Program Verifier*, in "6th Working Conference on Verified Software: Theories, Tools and Experiments (VSTTE)", Vienna, Austria, D. GIANNAKOPOULOU, D. KROENING (editors), Lecture Notes in Computer Science, Springer, July 2014, vol. 8471, p. 37–51, http://hal.inria.fr/hal-01067197.

[77] M. CLOCHARD, L. GONDELMAN.*Double WP: vers une preuve automatique d'un compilateur*, in "Vingt-sixièmes Journées Francophones des Langages Applicatifs", Val d'Ajol, France, January 2015, https://hal.inria.fr/hal-01094488.

[78] M. CLOCHARD, L. GONDELMAN, M. PEREIRA.*The Matrix Reproved*, in "8th Working Conference on Verified Software: Theories, Tools and Experiments (VSTTE)", Toronto, Canada, S. BLAZY, M. CHECHIK (editors), Lecture Notes in Computer Science, Springer, July 2016, https://hal.inria.fr/hal-01316902.

[79] M. CLOCHARD, C. MARCHÉ, A. PASKEVICH.*Verified Programs with Binders*, in "Programming Languages meets Program Verification (PLPV)", ACM Press, 2014, http://hal.inria.fr/hal-00913431.

[80] S. CONCHON.*SMT Techniques and their Applications: from Alt-Ergo to Cubicle*, Université Paris-Sud, December 2012, In English, http://www.lri.fr/~conchon/publis/conchonHDR.pdf, Thèse d'habilitation.

[81] S. CONCHON, É. CONTEJEAN, M. IGUERNELALA.*Canonized Rewriting and Ground AC Completion Modulo Shostak Theories*, in "Tools and Algorithms for the Construction and Analysis of Systems", Saarbrücken, Germany, P. A. ABDULLA, K. R. M. LEINO (editors), Lecture Notes in Computer Science, Springer, April 2011, vol. 6605, p. 45-59, http://hal.inria.fr/hal-00777663.

[82] S. CONCHON, É. CONTEJEAN, M. IGUERNELALA.*Canonized Rewriting and Ground AC Completion Modulo Shostak Theories : Design and Implementation*, in "Logical Methods in Computer Science", September 2012, vol. 8, n^{o} 3, p. 1–29, http://hal.inria.fr/hal-00798082.

[83] S. CONCHON, D. DECLERCK, L. MARANGET, A. MEBSOUT.*Vérification de programmes C concurrents avec Cubicle : Enfoncer les barrières*, in "Vingt-cinquièmes Journées Francophones des Langages Applicatifs", Fréjus, France, January 2014, https://hal.inria.fr/hal-01088655.

[84] S. CONCHON, A. GOEL, S. KRSTIĆ, A. MEBSOUT, F. ZAÏDI.*Invariants for Finite Instances and Beyond*, in "FMCAD", Portland, Oregon, États-Unis, October 2013, p. 61–68, http://hal.archives-ouvertes.fr/hal-00924640.

[85] S. CONCHON, M. IGUERNELALA.*Tuning the Alt-Ergo SMT Solver for B Proof Obligations*, in "Abstract State Machines, Alloy, B, VDM, and Z (ABZ)", Toulouse, France, Lecture Notes in Computer Science, Springer, June 2014, vol. 8477, p. 294–297, https://hal.inria.fr/hal-01093000.

[86] S. CONCHON, M. IGUERNELALA, A. MEBSOUT.*A Collaborative Framework for Non-Linear Integer Arithmetic Reasoning in Alt-Ergo*, 2013, https://hal.archives-ouvertes.fr/hal-00924646.

[87] S. CONCHON, A. MEBSOUT, F. ZAÏDI.*Vérification de systèmes paramétrés avec Cubicle*, in "Vingt-quatrièmes Journées Francophones des Langages Applicatifs", Aussois, France, February 2013, http://hal.inria.fr/hal-00778832.

[88] S. CONCHON, G. MELQUIOND, C. ROUX, M. IGUERNELALA.*Built-in Treatment of an Axiomatic Floating-Point Theory for SMT Solvers*, in "SMT workshop", Manchester, UK, P. FONTAINE, A. GOEL (editors), LORIA, 2012, p. 12–21.

[89] M. DAHLWEID, M. MOSKAL, T. SANTEN, S. TOBIES, W. SCHULTE.*VCC: Contract-based modular verification of concurrent C*, in "31st International Conference on Software Engineering, ICSE 2009, May 16-24, 2009, Vancouver, Canada, Companion Volume", IEEE Comp. Soc. Press, 2009, p. 429-430.

[90] D. DELAHAYE, C. DUBOIS, C. MARCHÉ, D. MENTRÉ.*The BWare Project: Building a Proof Platform for the Automated Verification of B Proof Obligations*, in "Abstract State Machines, Alloy, B, VDM, and Z (ABZ)", Toulouse, France, Lecture Notes in Computer Science, Springer, June 2014, vol. 8477, p. 290–293, http://hal.inria.fr/hal-00998092/en/.

[91] D. DELAHAYE, C. MARCHÉ, D. MENTRÉ.*Le projet BWare : une plate-forme pour la vérification automatique d'obligations de preuve B*, in "Approches Formelles dans l'Assistance au Développement de Logiciels (AFADL)", Paris, France, EasyChair, June 2014, http://hal.inria.fr/hal-00998094/en/.

[92] C. DROSS, S. CONCHON, J. KANIG, A. PASKEVICH.*Reasoning with Triggers*, in "SMT workshop", Manchester, UK, P. FONTAINE, A. GOEL (editors), LORIA, 2012.

[93] C. DROSS.*Generic Decision Procedures for Axiomatic First-Order Theories*, Université Paris-Sud, April 2014, http://tel.archives-ouvertes.fr/tel-01002190.

[94] J.-C. FILLIÂTRE.*Combining Interactive and Automated Theorem Proving in Why3 (invited talk)*, in "Automation in Proof Assistants 2012", Tallinn, Estonia, K. HELJANKO, H. HERBELIN (editors), April 2012.

[95] J.-C. FILLIÂTRE.*Combining Interactive and Automated Theorem Proving using Why3 (invited tutorial)*, in "Second International Workshop on Intermediate Verification Languages (BOOGIE 2012)", Berkeley, California, USA, Z. RAKAMARIĆ (editor), July 2012.

[96] J.-C. FILLIÂTRE.*Verifying Two Lines of C with Why3: an Exercise in Program Verification*, in "Verified Software: Theories, Tools, Experiments (4th International Conference VSTTE)", Philadelphia, USA, R. JOSHI, P. MÜLLER, A. PODELSKI (editors), Lecture Notes in Computer Science, Springer, January 2012, vol. 7152, p. 83–97.

[97] J.-C. FILLIÂTRE.*Deductive Program Verification*, in "Programming Languages Mentoring Workshop (PLMW 2013)", Rome, Italy, N. FOSTER, P. GARDNER, A. SCHMITT, G. SMITH, P. THIEMAN, T. WRIGSTAD (editors), January 2013, http://hal.inria.fr/hal-00799190.

[98] J.-C. FILLIÂTRE.*One Logic To Use Them All*, in "24th International Conference on Automated Deduction (CADE-24)", Lake Placid, USA, Lecture Notes in Artificial Intelligence, Springer, June 2013, vol. 7898, p. 1–20, http://hal.inria.fr/hal-00809651/en/.

[99] J.-C. FILLIÂTRE, A. PASKEVICH.*Why3 — Where Programs Meet Provers*, in "Proceedings of the 22nd European Symposium on Programming", M. FELLEISEN, P. GARDNER (editors), Lecture Notes in Computer Science, Springer, March 2013, vol. 7792, p. 125–128, http://hal.inria.fr/hal-00789533.

[100] J.-C. FILLIÂTRE, A. PASKEVICH, A. STUMP.*The 2nd Verified Software Competition: Experience Report*, in "COMPARE2012: 1st International Workshop on Comparative Empirical Evaluation of Reasoning Systems", Manchester, UK, V. KLEBANOV, S. GREBING (editors), EasyChair, June 2012, http://hal.inria.fr/hal-00798777.

[101] P. HERMS.*Certification of a Tool Chain for Deductive Program Verification*, Université Paris-Sud, January 2013, http://tel.archives-ouvertes.fr/tel-00789543.

[102] P. HERMS, C. MARCHÉ, B. MONATE.*A Certified Multi-prover Verification Condition Generator*, in "Verified Software: Theories, Tools, Experiments (4th International Conference VSTTE)", Philadelphia, USA,

R. JOSHI, P. MÜLLER, A. PODELSKI (editors), Lecture Notes in Computer Science, Springer, January 2012, vol. 7152, p. 2–17, http://hal.inria.fr/hal-00639977.

[103] M. IGUERNELALA.*Strengthening the Heart of an SMT-Solver: Design and Implementation of Efficient Decision Procedures*, Université Paris-Sud, June 2013, http://tel.archives-ouvertes.fr/tel-00842555.

[104] D. ISHII, G. MELQUIOND, S. NAKAJIMA.*Inductive Verification of Hybrid Automata with Strongest Post-condition Calculus*, in "Proceedings of the 10th Conference on Integrated Formal Methods", Turku, Finland, E. B. JOHNSEN, L. PETRE (editors), Lecture Notes in Computer Science, 2013, vol. 7940, p. 139–153, http://hal.inria.fr/hal-00806701.

[105] J. KANIG, E. SCHONBERG, C. DROSS.*Hi-Lite: the convergence of compiler technology and program verification*, in "Proceedings of the 2012 ACM Conference on High Integrity Language Technology, HILT '12", Boston, USA, B. BROSGOL, J. BOLENG, S. T. TAFT (editors), ACM Press, 2012, p. 27–34.

[106] G. KLEIN, J. ANDRONICK, K. ELPHINSTONE, G. HEISER, D. COCK, P. DERRIN, D. ELKADUWE, K. ENGELHARDT, R. KOLANSKI, M. NORRISH, T. SEWELL, H. TUCH, S. WINWOOD.*seL4: Formal verification of an OS kernel*, in "Communications of the ACM", June 2010, vol. 53, n^o 6, p. 107–115.

[107] C. LELAY.*A New Formalization of Power Series in Coq*, in "5th Coq Workshop", Rennes, France, July 2013, p. 1–2, http://hal.inria.fr/hal-00880212.

[108] C. LELAY.*Coq passe le bac*, in "JFLA - Journées francophones des langages applicatifs", Fréjus, France, January 2014.

[109] C. LELAY, G. MELQUIOND.*Différentiabilité et intégrabilité en Coq. Application à la formule de d'Alembert*, in "Vingt-troisièmes Journées Francophones des Langages Applicatifs", Carnac, France, February 2012, http://hal.inria.fr/hal-00642206/fr/.

[110] X. LEROY.*A formally verified compiler back-end*, in "Journal of Automated Reasoning", 2009, vol. 43, n^o 4, p. 363–446, http://hal.inria.fr/inria-00360768/en/.

[111] A. MAHBOUBI, G. MELQUIOND, T. SIBUT-PINOTE.*Formally Verified Approximations of Definite Integrals*, in "Proceedings of the 7th International Conference on Interactive Theorem Proving", Nancy, France, J. C. BLANCHETTE, S. MERZ (editors), Lecture Notes in Computer Science, August 2016, vol. 9807, https://hal.inria.fr/hal-01289616.

[112] C. MARCHÉ, A. TAFAT.*Weakest Precondition Calculus, revisited using Why3*, Inria, December 2012, n^o RR-8185, http://hal.inria.fr/hal-00766171.

[113] C. MARCHÉ, A. TAFAT.*Calcul de plus faible précondition, revisité en Why3*, in "Vingt-quatrièmes Journées Francophones des Langages Applicatifs", Aussois, France, February 2013, http://hal.inria.fr/hal-00778791.

[114] C. MARCHÉ.*Verification of the Functional Behavior of a Floating-Point Program: an Industrial Case Study*, in "Science of Computer Programming", March 2014, vol. 96, n^o 3, p. 279–296, http://hal.inria.fr/hal-00967124/en.

[115] É. MARTIN-DOREL, G. MELQUIOND, J.-M. MULLER.*Some Issues related to Double Roundings*, in "BIT Numerical Mathematics", 2013, vol. 53, n^{o} 4, p. 897–924, http://hal-ens-lyon.archives-ouvertes.fr/ensl-00644408.

[116] A. MEBSOUT.*Invariants inference for model checking of parameterized systems*, Université Paris-Sud, September 2014, https://tel.archives-ouvertes.fr/tel-01073980.

[117] G. MELQUIOND.*Floating-point arithmetic in the Coq system*, in "Information and Computation", 2012, vol. 216, p. 14–23, http://hal.inria.fr/hal-00797913.

[118] G. MELQUIOND, W. G. NOWAK, P. ZIMMERMANN.*Numerical Approximation of the Masser-Gramain Constant to Four Decimal Digits: delta=1.819...*, in "Mathematics of Computation", 2013, vol. 82, p. 1235–1246, http://hal.inria.fr/hal-00644166/en/.

[119] D. MENTRÉ, C. MARCHÉ, J.-C. FILLIÂTRE, M. ASUKA.*Discharging Proof Obligations from Atelier B using Multiple Automated Provers*, in "ABZ'2012 - 3rd International Conference on Abstract State Machines, Alloy, B and Z", Pisa, Italy, S. REEVES, E. RICCOBENE (editors), Lecture Notes in Computer Science, Springer, June 2012, vol. 7316, p. 238–251, http://hal.inria.fr/hal-00681781/en/.

[120] J.-M. MULLER, N. BRISEBARRE, F. DE DINECHIN, C.-P. JEANNEROD, V. LEFÈVRE, G. MELQUIOND, N. REVOL, D. STEHLÉ, S. TORRES.*Handbook of Floating-Point Arithmetic*, Birkhäuser, 2010.

[121] T. M. T. NGUYEN, C. MARCHÉ.*Hardware-Dependent Proofs of Numerical Programs*, in "Certified Programs and Proofs", J.-P. JOUANNAUD, Z. SHAO (editors), Lecture Notes in Computer Science, Springer, December 2011, p. 314–329, http://hal.inria.fr/hal-00772508.

[122] T. M. T. NGUYEN.*Taking architecture and compiler into account in formal proofs of numerical programs*, Université Paris-Sud, June 2012, http://tel.archives-ouvertes.fr/tel-00710193.

[123] M. NORRISH.*C Formalised in HOL*, University of Cambridge, November 1998.

[124] M. PEREIRA, J.-C. FILLIÂTRE, S. M. DE SOUSA.*ARMY: a Deductive Verification Platform for ARM Programs Using Why3*, in "INForum 2012", September 2012.

[125] P. ROUX.*Formal Proofs of Rounding Error Bounds*, in "Journal of Automated Reasoning", 2015, https://hal.archives-ouvertes.fr/hal-01091189.

[126] N. SCHIRMER.*Verification of Sequential Imperative Programs in Isabelle/HOL*, Technische Universität München, 2006.

[127] A. TAFAT.*Preuves par raffinement de programmes avec pointeurs*, Université Paris-Sud, September 2013, http://tel.archives-ouvertes.fr/tel-00874679.

[128] F. DE DINECHIN, C. LAUTER, G. MELQUIOND.*Certifying the floating-point implementation of an elementary function using Gappa*, in "IEEE Transactions on Computers", 2011, vol. 60, n^{o} 2, p. 242–253, http://hal.inria.fr/inria-00533968/en/.

[129] L. DE MOURA, N. BJØRNER.*Z3, An Efficient SMT Solver*, in "TACAS", Lecture Notes in Computer Science, Springer, 2008, vol. 4963, p. 337–340.

Team TROPICAL

TROPICAL

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions).

RESEARCH CENTER
Saclay - Île-de-France

THEME
Optimization and control of dynamic systems

Table of contents

Team TROPICAL

Creation of the Team: 2016 January 01

Keywords:

Computer Science and Digital Science:

A2.4. - Verification, reliability, certification
A6.2.5. - Numerical Linear Algebra
A6.2.6. - Optimization
A6.4.1. - Deterministic control
A6.4.2. - Stochastic control
A8.1. - Discrete mathematics, combinatorics
A8.2. - Optimization
A8.3. - Geometry, Topology
A8.9. - Performance evaluation
A8.11. - Game Theory

Other Research Topics and Application Domains:

B1.1.10. - Mathematical biology
B4.4. - Energy delivery
B4.4.1. - Smart grids
B9.9. - Risk management

1. Personnel

Research Scientists

Stéphane Gaubert [Team leader, Inria, Senior Researcher, HDR]
Marianne Akian [Inria, Senior Researcher, HDR]
Xavier Allamigeon [Corps des Mines, under secondment, Inria, Senior Researcher]
Cormac Walsh [Inria, Researcher]

Technical Staff

Vasileios Charisopoulos [Inria, from May 2017 until Aug 2017]

PhD Students

Vianney Boeuf [Ingénieur du corps des Ponts, ENPC, PhD Student, until Aug 2017]
Jean-Bernard Eytard [Inria, PhD Student]
Eric Fodjo [I-Fihn Consulting, Consultant, PhD Student]
Paulin Jacquot [EDF (Cifre), PhD Student]
Mateusz Skomra [Ecole polytechnique, PhD Student]
Nikolas Stott [Inria, PhD Student]
Duy Nghi, Benoît Tran [Univ Paris-Est Marne La Vallée, PhD Student, from Sep 2017]

Post-Doctoral Fellows

Daniel Jones [Inria, Post-Doctoral Fellow, from Aug 2017]
Marie Maccaig [FMJH, Ecole polytechnique, Post-Doctoral Fellow, until Mar 2017]
Aurelien Sagnier [Ecole polytechnique, from Sep 2017]

Visiting Scientists

Zheng Hua [HKU, from Jun 2017 until Jul 2017]

Ricardo Katz [CONICET, from May 2017 until Jul 2017]
Gleb Koshevoy [Russian Academy of Sciences, from Feb 2017 until Mar 2017]
Zheng Qu [HKU, from May 2017 until Jul 2017]

Administrative Assistants
Hanadi Dib [Inria, from Sep 2017]
Corinne Petitot [Inria, until Sep 2017]

2. Overall Objectives

2.1. Introduction

The project develops tropical methods motivated by applications arising in decision theory (deterministic and stochastic optimal control, game theory, optimization and operations research), in the analysis or control of classes of dynamical systems (including timed discrete event systems and positive systems), in the verification of programs and systems, and in the development of numerical algorithms. Tropical algebra tools are used in interaction with various methods, coming from convex analysis, Hamilton–Jacobi partial differential equations, metric geometry, Perron-Frobenius and nonlinear fixed-point theories, combinatorics or algorithmic complexity. The emphasis of the project is on mathematical modelling and computational aspects.

The subtitle of the *Tropical* project, namely, “structures, algorithms, and interactions”, refers to the spirit of our research, including a methodological component, computational aspects, and finally interactions with other scientific fields or real world applications, in particular through mathematical modelling.

2.2. Scientific context

Tropical algebra, geometry, and analysis have enjoyed spectacular development in recent years. Tropical structures initially arose to solve problems in performance evaluation of discrete event systems [61], combinatorial optimization [64], or automata theory [105]. They also arose in mathematical physics and asymptotic analysis [95], [92]. More recently, these structures have appeared in several areas of pure mathematics, in particular in the study of combinatorial aspects of algebraic geometry [84], [117], [107], [90], in algebraic combinatorics [78], and in arithmetics [68]. Also, further applications of tropical methods have appeared, including optimal control [99], program invariant computation [55] and timed systems verification [94], and zero-sum games [2].

The term ‘tropical’ generally refers to algebraic structures in which the laws originate from optimization processes. The prototypical tropical structure is the max-plus semifield, consisting of the real numbers, equipped with the maximum, thought of as an additive law, and the addition, thought of as a multiplicative law. Tropical objects appear as limits of classical objects along certain deformations (“log-limits sets” of Bergman, “Maslov dequantization”, or “Viro deformation”). For this reason, the introduction of tropical tools often yields new insights into old familiar problems, leading either to counterexamples or to new methods and results; see for instance [117], [101]. In some applications, like optimal control, discrete event systems, or static analysis of programs, tropical objects do not appear through a limit procedure, but more directly as a modelling or computation/analysis tool; see for instance [112], [61], [86], [65].

Tropical methods are linked to the fields of positive systems and of metric geometry [103], [11]. Indeed, tropically linear maps are monotone (a.k.a. order-preserving). They are also nonexpansive in certain natural metrics (sup-norm, Hopf oscillation, Hilbert’s projective metric, ...). In this way, tropical dynamical systems appear to be special cases of nonexpansive, positive, or monotone dynamical systems, which are studied as part of linear and non-linear Perron-Frobenius theory [93], [3]. Such dynamical systems are of fundamental importance in the study of repeated games [100]. Monotonicity properties are also essential in the understanding of the fixed points problems which determine program invariants by abstract interpretation [69]. The latter problems are actually somehow similar to the ones arising in the study of zero-sum games; see [7]. Moreover, positivity or monotonicity methods are useful in population dynamics, either in a discrete space setting [114] or in a PDE setting [62]. In such cases, solving tropical problems often leads to solutions or combinatorial insights on classical problems involving positivity conditions (e.g., finding equilibria of dynamical systems with nonnegative coordinates, understanding the qualitative and quantitative behavior of growth rates / Floquet eigenvalues [9], etc). Other applications of Perron-Frobenius theory originate from quantum information and control [106], [111].

3. Research Program

3.1. Optimal control and zero-sum games

The dynamic programming approach allows one to analyze one or two-player dynamic decision problems by means of operators, or partial differential equations (Hamilton–Jacobi or Isaacs PDEs), describing the time evolution of the value function, i.e., of the optimal reward of one player, thought of as a function of the initial state and of the horizon. We work especially with problems having long or infinite horizon, modelled by stopping problems, or ergodic problems in which one optimizes a mean payoff per time unit. The determination of optimal strategies reduces to solving nonlinear fixed point equations, which are obtained either directly from discrete models, or after a discretization of a PDE.

The geometry of solutions of optimal control and game problems Basic questions include, especially for stationary or ergodic problems, the understanding of existence and uniqueness conditions for the solutions of dynamic programming equations, for instance in terms of controllability or ergodicity properties, and more generally the understanding of the structure of the full set of solutions of stationary Hamilton–Jacobi PDEs and of the set of optimal strategies. These issues are already challenging in the one-player deterministic case, which is an application of choice of tropical methods, since the Lax-Oleinik semigroup, i.e., the evolution semigroup of the Hamilton-Jacobi PDE, is a linear operator in the tropical sense. Recent progress in the deterministic case has been made by combining dynamical systems and PDE techniques (weak KAM theory [75]), and also using metric geometry ideas (abstract boundaries can be used to represent the sets of solutions [89], [4]). The two player case is challenging, owing to the lack of compactness of the analogue of the Lax-Oleinik semigroup and to a richer geometry. The conditions of solvability of ergodic problems for games (for instance, solvability of ergodic Isaacs PDEs), and the representation of solutions are only understood in special cases, for instance in the finite state space case, through tropical geometry and non-linear Perron-Frobenius methods [48],[12], [3].

Algorithmic aspects: from combinatorial algorithms to the attenuation of the curse of dimensionality Our general goal is to push the limits of solvable models by means of fast algorithms adapted to large scale instances. Such instances arise from discrete problems, in which the state space may so large that it is only accessible through local oracles (for instance, in some web ranking applications, the number of states may be the number of web pages) [76]. They also arise from the discretization of PDEs, in which the number of states grows exponentially with the number of degrees of freedom, according to the "curse of dimensionality". A first line of research is the development of *new approximation methods for the value function*. So far, classical approximations by linear combinations have been used, as well as approximation by suprema of linear or quadratic forms, which have been introduced in the setting of dual dynamic programming and of the so called "max-plus basis methods" [77]. We believe that more concise or more accurate approximations may be obtained by unifying these methods. Also, some max-plus basis methods have been shown to *attenuate the curse of dimensionality* for very special problems (for instance involving switching) [96], [80]. This suggests that the complexity of control or games problems may be measured by more subtle quantities that the mere number of states, for instance, by some forms of metric entropy (for example, certain large scale problems have a low complexity owing to the presence of decomposition properties, "highway hierarchies", etc.). A second line of of our research is the development of *combinatorial algorithms*, to solve large scale zero-sum two-player problems with discrete state space. This is related to current open problems in algorithmic game theory. In particular, the existence of polynomial-time algorithms for games with ergodic payment is an open question. See e.g. [5] for a polynomial time average complexity result derived by tropical methods. The two lines of research are related, as the understanding of the geometry of solutions allows to develop better approximation or combinatorial algorithms.

3.2. Non-linear Perron-Frobenius theory, nonexpansive mappings and metric geometry

Several applications (including population dynamics [9] and discrete event systems [61], [67], [54]) lead to studying classes of dynamical systems with remarkable properties: preserving a cone, preserving an order, or being nonexpansive in a metric. These can be studied by techniques of non-linear Perron-Frobenius theory [3] or metric geometry [10]. Basic issues concern the existence and computation of the "escape rate" (which determines the throughput, the growth rate of the population), the characterizations of stationary regimes (non-linear fixed points), or the study of the dynamical properties (convergence to periodic orbits). Nonexpansive mappings also play a key role in the "operator approach" to zero-sum games, since the one-day operators of games are nonexpansive in several metrics, see [8].

3.3. Tropical algebra and convex geometry

The different applications mentioned in the other sections lead us to develop some basic research on tropical algebraic structures and in convex and discrete geometry, looking at objects or problems with a "piecewise-linear " structure. These include the geometry and algorithmics of tropical convex sets [56], [50], tropical semialgebraic sets [59], the study of semi-modules (analogues of vector spaces when the base field is replaced by a semi-field), the study of systems of equations linear in the tropical sense, investigating for instance the analogues of the notions of rank, the analogue of the eigenproblems [14], and more generally of systems of tropical polynomial equations. Our research also builds on, and concern, classical convex and discrete geometry methods.

3.4. Tropical methods applied to optimization, perturbation theory and matrix analysis

Tropical algebraic objects appear as a deformation of classical objects thought various asymptotic procedures. A familiar example is the rule of asymptotic calculus,

$$e^{-a/\epsilon} + e^{-b/\epsilon} \asymp e^{-\min(a,b)/\epsilon} \;, \qquad e^{-a/\epsilon} \times e^{-b/\epsilon} = e^{-(a+b)/\epsilon} \;, \tag{16}$$

when $\epsilon \to 0^+$. Deformations of this kind have been studied in different contexts: large deviations, zero-temperature limits, Maslov's "dequantization method" [95], non-archimedean valuations, log-limit sets and Viro's patchworking method [117], etc.

This entails a relation between classical algorithmic problems and tropical algorithmic problems, one may first solve the $\epsilon = 0$ case (non-archimedean problem), which is sometimes easier, and then use the information gotten in this way to solve the $\epsilon = 1$ (archimedean) case.

In particular, tropicalization establishes a connection between polynomial systems and piecewise affine systems that are somehow similar to the ones arising in game problems. It allows one to transfer results from the world of combinatorics to "classical" equations solving. We investigate the consequences of this correspondence on complexity and numerical issues. For instance, combinatorial problems can be solved in a robust way. Hence, situations in which the tropicalization is faithful lead to improved algorithms for classical problems. In particular, scalings for the polynomial eigenproblems based on tropical preprocessings have started to be used in matrix analysis [82], [85].

Moreover, the tropical approach has been recently applied to construct examples of linear programs in which the central path has an unexpectedly high total curvature [53], and it has also led to positive polynomial-time average case results concerning the complexity of mean payoff games. Similarly, we are studying semidefinite programming over non-archimedean fields [59], [58], with the goal to better understand complexity issues in classical semidefinite and semi-algebraic programming.

4. Application Domains

4.1. Discrete event systems (manufacturing systems, networks)

One important class of applications of max-plus algebra comes from discrete event dynamical systems [61]. In particular, modelling timed systems subject to synchronization and concurrency phenomena leads to studying dynamical systems that are non-smooth, but which have remarkable structural properties (nonexpansiveness in certain metrics , monotonicity) or combinatorial properties. Algebraic methods allow one to obtain analytical expressions for performance measures (throughput, waiting time, etc). A recent application, to emergency call centers, can be found in [54].

4.2. Optimal control and games

Optimal control and game theory have numerous well established applications fields: mathematical economy and finance, stock optimization, optimization of networks, decision making, etc. In most of these applications, one needs either to derive analytical or qualitative properties of solutions, or design exact or approximation algorithms adapted to large scale problems.

4.3. Operations Research

We develop, or have developed, several aspects of operations research, including the application of stochastic control to optimal pricing, optimal measurement in networks [108]. Applications of tropical methods arise in particular from discrete optimization [65], [66], scheduling problems with and-or constraints [102], or product mix auctions [116].

4.4. Computing program and dynamical systems invariants

A number of programs and systems verification questions, in which safety considerations are involved, reduce to computing invariant subsets of dynamical systems. This approach appears in various guises in computer science, for instance in static analysis of program by abstract interpretation, along the lines of P. and R. Cousot [69], but also in control (eg, computing safety regions by solving Isaacs PDEs). These invariant sets are often sought in some tractable effective class: ellipsoids, polyhedra, parametric classes of polyhedra with a controlled complexity (the so called "templates" introduced by Sankaranarayanan, Sipma and Manna [109]), shadows of sets represented by linear matrix inequalities, disjunctive constraints represented by tropical polyhedra [55], etc. The computation of invariants boils down to solving large scale fixed point problems. The latter are of the same nature as the ones encountered in the theory of zero-sum games, and so, the techniques developed in the previous research directions (especially methods of monotonicity, nonexpansiveness, discretization of PDEs, etc) apply to the present setting, see e.g. [79], [83] for the application of policy iteration type algorithms, or for the application for fixed point problems over the space of quadratic forms [7]. The problem of computation of invariants is indeed a key issue needing the methods of several fields: convex and nonconvex programming, semidefinite programming and symbolic computation (to handle semialgebraic invariants), nonlinear fixed point theory, approximation theory, tropical methods (to handle disjunctions), and formal proof (to certify numerical invariants or inequalities).

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Performance evaluation of the 17-18-112 call center in Paris

Vianney Bœuf completed his PhD, done in collaboration with Brigade des Sapeurs Pompiers de Paris, on the performance evaluation of the new organization of the Paris emergency call center developed by Préfecture de Police. See Section 7.5.2.

5.1.2. Maximal upper bounds in Löwner order

A classical theorem of Kadison (1951) shows that the set of real quadratic forms, equipped with the pointwise order, is an antilattice, meaning that two quadratic forms have a least upper bound (or dually, a greatest lower bound) if and only if they are comparable. In [23], Nikolas Stott gave a quantitative version of Kadison theorem, characterizing the set of minimal upper bound as the quotient an indefinite orthogonal group. Applications of these ideas to hybrid systems verification appeared in [16], [30].

5.1.3. Formal proofs in linear programming

Xavier Allamigeon and Ricardo Katz have formalized in the proof assistant Coq several basic results in the theory of convex polyhedra and linear optimization. These include Farkas Lemma, the duality theorem of linear programming, separation from convex hulls, Minkowski Theorem, etc. See [27] and Section 7.3.1.

6. New Software and Platforms

6.1. Coq-Polyhedra

KEYWORDS: Coq - Polyhedra - Automated theorem proving - Linear optimization
SCIENTIFIC DESCRIPTION: Coq-Polyhedra is a library providing a formalization of convex polyhedra in the Coq proof assistant. While still in active development, it provides an implementation of the simplex method, and already handles the basic properties of polyhedra such as emptiness, boundedness, membership. Several fundamental results in the theory of convex polyhedra, such as Farkas Lemma, duality theorem of linear programming, and Minkowski Theorem, are also formally proved.

The formalization is based on the Mathematical Components library, and makes an extensive use of the boolean reflection methodology.
FUNCTIONAL DESCRIPTION: Coq-Polyhedra is a library which aims at formalizing convex polyhedra in Coq

- Participants: Xavier Allamigeon, Vasileios Charisopoulos and Ricardo Katz
- Partner: CIFASIS
- Contact: Xavier Allamigeon
- Publication: A Formalization of Convex Polyhedra Based on the Simplex Method
- URL: https://github.com/nhojem/Coq-Polyhedra

7. New Results

7.1. Optimal control and zero-sum games

7.1.1. Fixed points of order preserving homogeneous maps and zero-sum games

Participants: Marianne Akian, Stéphane Gaubert.

The PhD work of Antoine Hochart [88] was dealing with the applications of methods of non-linear fixed point theory to zero-sum games.

A highlight of his PhD is the characterization of the property of ergodicity for zero-sum games. In the special "zero-player" case, i.e., for a Markov chain equipped with an additive functional (payment) of the trajectory, the ergodicity condition entails that the mean payoff is independent of the initial state, for any choice of the payment. In the case of finite Markov chains, ergodicity admits several characterizations, including a combinatorial one (the uniqueness of the final class). This carries over to the two player case: ergodicity is now characterized by the absence of certain pairs of conjugate invariant sets (dominions), and it can be checked using directed hypergraphs algorithms. This leads to an explicit combinatorial sufficient condition for the solvability of the "ergodic equation", which is the main tool in the numerical approach of the mean payoff problem. These results appeared in [52] for the case of bounded paiements. A more general approach was developed in [87], in which zero-sum games are now studied abstractly in terms of accretive operators. This allows one to show that the bias vector (the solution of the ergodic equation) is unique for a generic perturbation of the payments. A more recent work include the introduction of an abstract game allowing us to deal with general monotone additively homogeneous operators and thus to unbounded paiements.

Another series of results of the thesis concern the finite action space, showing that the set of payments for which the bias vector is not unique coincides with the union of lower dimensional cells of a polyhedral complex, which an application to perturbation schemes in policy iteration [12].

A last result of the thesis is a representation theorem for "payment free" Shapley operators, showing that these are characterized by monotonicity and homogeneity axioms [13]. This extends to the two-player case known representation theorems for risk measures.

7.1.2. The operator approach to entropy games

Participants: Marianne Akian, Stéphane Gaubert.

Entropy games were recently introduced by Asarin et al. A player (Despot) wishes to minimize a measure of "freedom" given by a topological entropy, whereas the other player (Tribune) wishes to maximize it. In [25], we developed an operator approach for entropy games. We showed that they reduce to risk sensitive type game problems, and deduced that entropy games in Despot has a few positions with non-trivial actions can be solved in polynomial time.

7.1.3. Probabilistic and max-plus approximation of Hamilton-Jacobi-Bellman equations

Participants: Marianne Akian, Eric Fodjo.

The PhD thesis of Eric Fodjo concerns stochastic control problems obtained in particular in the modelisation of portfolio selection with transaction costs. The dynamic programming method leads to a Hamilton-Jacobi-Bellman partial differential equation, on a space with a dimension at least equal to the number of risky assets. The curse of dimensionality does not allow one to solve numerically these equations for a large dimension (greater to 5). We propose to tackle these problems with numerical methods combining policy iterations, probabilistic discretisations, max-plus discretisations, in order to increase the possible dimension.

We consider fully nonlinear Hamilton-Jacobi-Bellman equations associated to diffusion control problems with finite horizon involving a finite set-valued (or switching) control and possibly a continuum-valued control. In [47], we constructed a lower complexity probabilistic numerical algorithm by combining the idempotent expansion properties obtained by McEneaney, Kaise and Han [91], [97] for solving such problems with a numerical probabilistic method such as the one proposed by Fahim, Touzi and Warin [74] for solving some fully nonlinear parabolic partial differential equations, when the volatility does not oscillate too much. In [38], [39] (also presented in [24]), we improve the method of Fahim, Touzi and Warin by introducing probabilistic schemes which are monotone without any restrictive condition, allowing one to solve fully nonlinear parabolic partial differential equations with general volatilities. We study the convergence and obtain error estimates when the parameters and the value function are bounded. We are now studying the more general quadratic growth case.

7.1.4. Tropical-SDDP algorithms for stochastic control problems involving a switching control

Participants: Marianne Akian, Duy Nghi, Benoît Tran.

The PhD thesis of Benoît Tran, supervised by Jean-Philippe Chancelier (ENPC) and Marianne Akian concerns the numerical solution of the dynamic programming equation of discrete time stochastic control problems.

Several methods have been proposed in the litterature to bypass the curse of dimensionality difficulty of such an equation, by assuming a certain structure of the problem. Examples are the max-plus based method of McEneaney [98], [99], the stochastic dual dynamic programming (SDDP) algorithm of Pereira and Pinto [104], the mixed integer dynamic approximation scheme of Philpott, Faisal and Bonnans [60], the probabilistic numerical method of Fahim, Touzi and Warin [74]. We propose to associate and compare these methods in order to solve more general structures, in particular problems involving a finite set-valued (or switching) control and a continuum-valued control, with the property that the value function associated to a fixed switching strategy is convex.

7.2. Non-linear Perron-Frobenius theory, nonexpansive mappings and metric geometry

7.2.1. Order reversing maps on cones

Participant: Cormac Walsh.

We have been studying non-linear operators on open cones, particularly ones that preserve or reverse the order structure associated to the cone. A bijective map that preserves the order in both directions is called an order isomorphism. Those that reverse the order in both directions are order antimorphisms. These are closely related to the isometries of the Hilbert and Thompson metrics on the cone.

Previously, we have shown [118] that if there exists an antimorphism on an open cone that is homogeneous of degree -1, then the cone must be a symmetric cone, that is, have a transitive group of linear automorphisms and be self-dual.

The technique was to consider the Funk metric on the cone, which is a non-symmetric metric defined using the order and homogeneity structures. Each antimorphism on a cone that is homogeneous of degree -1 reverses this metric, and so interchanges the two horofunction boundaries of the metric, the one in the forward direction and the one in the backward direction. By studying these boundaries we obtained the result.

More recently, we have shown [45] that the homogeneity assumption is not actually necessary: every antimorphism on a cone is automatically homogeneous of degree -1.

Without the homogeneity assumption, the metric techniques do not work. Instead, it was necessary to study how the map acts on line segments parallel to extreme rays of the cone. This is similar to the what was done by Rothaus, Noll and Schäffer, and Artstein-Avidan and Slomka in their work on order isomorphisms. This means that our proof is essentially finite dimensional. Indeed, there are many interesting cones in infinite dimension that have few or no extreme rays.

In infinite dimension, it is natural to consider order unit spaces as a generalisation of cones, and JB-algebras as a generalisation of symmetric cones. Lemmens, Roelands, and Wortel have asked whether the existence of an order antimorphism that is homogeneous of degree -1 on the cone of an order-unit space implies that the space is a JB-algebra? The result above suggests further that one might even be able to drop the homogeneity assumption.

7.2.2. The set of minimal upper bounds of two matrices in the Loewner order

Participant: Nikolas Stott.

A classical theorem of Kadison shows that the space of symmetric matrices equipped with the Loewner order is an anti-lattice, meaning that two matrices have a least upper bound if and only if they are comparable. In [115], we refined this theorem by characterizing the set of minimal upper bounds: we showed that it is homeomorphic to the quotient space $O(p) \setminus O(p,q)/O(q)$, where $O(p,q)$ denotes the orthogonal group associated to the quadratic form with signature (p,q), and $O(p)$ denotes the standard pth orthogonal group.

7.2.3. Checking the strict positivity of Kraus maps is NP-hard

Participant: Stéphane Gaubert.

In collaboration with Zheng Qu (now with HKU, Hong Kong), I studied several decision problems arising from the spectral theory of Kraus maps (trace preserving completely positive maps), acting on the cone of positive semidefinite matrices. The latter appear in quantum information. We showed that checking the irreducibility (absence of non-trivial invariant face of the cone) and primitivity properties (requiring the iterates of the map to send the cone to its interior) can be checked in polynomial time, whereas checking positivity (whether the map sends the cone to its interior) is NP-hard. In [20], we studied complexity issues related to Kraus maps, and showed in particular that checking whether a Kraus map sends the cone to its interior is NP-hard.

7.3. Tropical algebra and convex geometry

7.3.1. Formalizing convex polyhedra in Coq

Participants: Xavier Allamigeon, Ricardo Katz [Conicet, Argentine].

In [27], we have made the first steps of a formalization of the theory of convex polyhedra in the proof assistant Coq. The originality of our approach lies in the fact that our formalization is carried out in an effective way, in the sense that the basic predicates over polyhedra (emptiness, boundedness, membership, etc) are defined by means of Coq programs. All these predicates are then proven to correspond to the usual logical statements. The latter take the form of the existence of certificates: for instance, the emptiness of a polyhedron is shown to be equivalent to the existence of a certificate *a la* Farkas. This equivalence between Boolean predicates and formulas living in the kind `Prop` is implemented by using the boolean reflection methodology, and the supporting tools provided by the Mathematical Components library and its tactic language. The benefit of the effective nature of our approach is demonstrated by the fact that we easily arrive at the proof of important results on polyhedra, such as several versions of Farkas Lemma, duality theorem of linear programming, separation from convex hulls, Minkowski Theorem, etc.

Our effective approach is made possible by implementing the simplex method inside Coq, and proving its correctness and termination. Two difficulties need to be overcome to formalize it. On the one hand, we need to deal with its termination. More precisely, the simplex method iterates over the so-called bases. Its termination depends on the specification of a pivoting rule, whose aim is to determine, at each iteration, the next basis. In this work, we have focused on proving that the lexicographic rule ensures termination. On the other hand, the simplex method is actually composed of two parts. The part that we previously described, called Phase II, requires an initial basis to start with. Finding such a basis is the purpose of Phase I. It consists in building an extended problem (having a trivial initial basis), and applying to it Phase II. Both phases need to be formalized to obtain a fully functional algorithm. a

7.3.2. Tropical totally positive matrices

Participant: Stéphane Gaubert.

In [81] (joint work with Adi Niv) we investigate the tropical analogues of totally positive and totally non-negative matrices, i.e, the images by the valuation of the corresponding classes of matrices over a non-archimedean field. We show in particular that tropical totally positive matrices essentially coincide with the Monge matrices (defined by the positivity of 2×2 tropical minors), arising in optimal transport. More recent developments include relations between tropical total positivity and planar networks.

7.3.3. Tropical compound matrix identities

Participants: Marianne Akian, Stéphane Gaubert.

A number of polynomial identities in tropical semirings can be derived from their classical analogues by application of a transfer principle [49], [51]. In the present work [40], joint with Adi Niv, we prove identities on compound matrices in extended tropical semirings, which cannot be obtained by transfer principles, but are rather obtained by combinatorial methods. Such identities include analogues to properties of conjugate matrices, powers of matrices and $\mathrm{adj}\,(A)\det\,(A)^{-1}$, all of which have implications on the eigenvalues of the corresponding matrices. A tropical Sylvester-Franke identity is provided as well.

7.3.4. Group algebra in characteristic one and invariant distances over finite groups

Participant: Stéphane Gaubert.

In [19] (joint work with Dominique Castella), we investigated a tropical analogue of group algebras. We studied tropical characters and related them to invariant distances over groups.

7.3.5. Volume and integer points of tropical polytopes

Participants: Marie Maccaig, Stéphane Gaubert.

We investigate in [43] the volume of tropical polytopes, as well as the number of integer points contained in integer polytopes. We proved that even approximating these values for a tropical polytope given by its vertices is hard, with no approximation algorithm with factor $2^{\mathrm{poly}(m,n)}$ existing unless $\mathrm{P} = \mathrm{NP}$. We further proved the $\sharp$P-hardness for the analogous problems for tropical polytopes instead defined by inequalities.

7.4. Tropical methods applied to optimization, perturbation theory and matrix analysis

7.4.1. Majorization inequalities for valuations of eigenvalues using tropical algebra

Participants: Marianne Akian, Stéphane Gaubert.

In [14], with Meisam Sharify (IPM, Tehran, Iran), we establish log-majorization inequalities of the eigenvalues of matrix polynomials using the tropical roots of some scalar polynomials depending only on the norms of the matrix coefficients. This extends to the case of matrix polynomials some bounds obtained by Hadamard, Ostrowski and Pólya for the roots of scalar polynomials.

These works have been presented in [46].

7.4.2. Tropicalization of the central path and application to the complexity of interior point methods

Participants: Xavier Allamigeon, Stéphane Gaubert.

This work is in collaboration with Pascal Benchimol (EDF Labs) and Michael Joswig (TU Berlin).

In optimization, path-following interior point methods are driven to an optimal solution along a trajectory called the central path. The *central path* of a linear program $\mathrm{LP}(A,b,c) \equiv \min\{c\cdot x \mid Ax \le b,\ x \ge 0\}$ is defined as the set of the optimal solutions (x^μ, w^μ) of the barrier problems:

$$\begin{array}{ll} \text{minimize} & c\cdot x - \mu(\sum_{j=1}^{n} \log x_j + \sum_{i=1}^{m} \log w_i) \\ \text{subject to} & Ax + w = b,\ x > 0,\ w > 0 \end{array}$$

While the complexity of interior point methods is known to be polynomial, an important question is to study the number of iterations which are performed by interior point methods, in particular whether it can be bounded by a polynomial in the dimension (mn) of the problem. This is motivated by Smale 9th problem [113], on the existence of a strongly polynomial complexity algorithm for linear programming. So far, this question has been essentially addressed though the study of the curvature of the central path, which measures how far a path differs from a straight line, see [71], [70], [73], [72]. In particular, by analogy with the classical Hirsch conjecture, Deza, Terlaky and Zinchencko [72] proposed the "continuous analogue of the Hirsch conjecture", which says that the total curvature of the central path is linearly bounded in the number m of constraints.

In a work of X. Allamigeon, P. Benchimol, S. Gaubert, and M. Joswig [41], we prove that primal-dual log-barrier interior point methods are not strongly polynomial, by constructing a family of linear programs with $3r+1$ inequalities in dimension $2r$ for which the number of iterations performed is in $\Omega(2^r)$. The total curvature of the central path of these linear programs is also exponential in r, disproving the continuous analogue of the Hirsch conjecture.

Our method is to tropicalize the central path in linear programming. The tropical central path is the piecewise-linear limit of the central paths of parameterized families of classical linear programs viewed through logarithmic glasses. We give an explicit geometric characterization of the tropical central path, as a tropical analogue of the barycenter of a sublevel set of the feasible set induced by the duality gap. We study the convergence properties of the classical central path to the tropical one. This allows us to show that that the number of iterations performed by interior point methods is bounded from below by the number of tropical segments constituting the tropical central path.

7.4.3. Tropical approach to semidefinite programming

Participants: Xavier Allamigeon, Stéphane Gaubert, Mateusz Skomra.

Semidefinite programming consists in optimizing a linear function over a spectrahedron. The latter is a subset of $\mathbb{R}^n$ defined by linear matrix inequalities, i.e., a set of the form

$$\left\{x \in \mathbb{R}^n : Q^{(0)} + x_1 Q^{(1)} + \cdots + x_n Q^{(n)} \succeq 0\right\}$$

where the $Q^{(k)}$ are symmetric matrices of order m, and $\succeq$ denotes the Loewner order on the space of symmetric matrices. By definition, $X \succeq Y$ if and only if $X - Y$ is positive semidefinite.

Semidefinite programming is a fundamental tool in convex optimization. It is used to solve various applications from engineering sciences, and also to obtain approximate solutions or bounds for hard problems arising in combinatorial optimization and semialgebraic optimization.

A general issue in computational optimization is to develop combinatorial algorithms for semidefinite programming. Indeed, semidefinite programs are usually solved via interior point methods. However, the latter provide an approximate solution in a polynomial number of iterations, provided that a strictly feasible initial solution. Semidefinite programming becomes a much harder matter if one requires an exact solution. The feasibility problem belongs to $\mathsf{NP}_{\mathbb{R}} \cap \mathsf{coNP}_{\mathbb{R}}$, where the subscript $\mathbb{R}$ refers to the BSS model of computation. It is not known to be in NP in the bit model.

We address semidefinite programming in the case where the field $\mathbb{R}$ is replaced by a nonarchimedean field, like the field of Puiseux series. In this case, methods from tropical geometry can be applied and are expected to allow one, in generic situations, to reduce semialgebraic problems to combinatorial problems, involving only the nonarchimedean valuations (leading exponents) of the coefficients of the input.

To this purpose, we first study tropical spectrahedra, which are defined as the images by the valuation of nonarchimedean spectrahedra. We establish that they are closed semilinear sets, and that, under a genericity condition, they are described by explicit inequalities expressing the nonnegativity of tropical minors of order 1 and 2. These results are gathered in the preprint [59].

Then, we show in [17] that the feasibility problem for a generic tropical spectrahedron is equivalent to solving a stochastic mean payoff game (with perfect information). The complexity of these games is a long-standing open problem. They are not known to be polynomial, however they belong to the class $\mathsf{NP} \cap \mathsf{coNP}$, and they can be solved efficiently in practice. This allows to apply stochastic game algorithms to solve nonarchimedean semidefinite feasibility problems. We obtain in this way both theoretical bounds and a practicable method which solves some large scale instances.

A long-standing problem is to characterize the convex semialgebraic sets that are SDP representable, meaning that they can be represented as the image of a spectrahedron by a (linear) projector. Helton and Nie conjectured that every convex semialgebraic set over the field of real numbers are SDP representable. Recently, [110] disproved this conjecture. In [26], we show, however, that the following result, which may be thought of as a tropical analogue of this conjecture, is true: over a real closed nonarchimedean field of Puiseux series, the convex semialgebraic sets and the projections of spectrahedra have precisely the same images by the nonarchimedean valuation. The proof relies on game theory methods applied to our previous results [59] and [17].

7.5. Applications

7.5.1. Geometry of the Loewner order and application to the synthesis of quadratic invariants in static analysis of program

Participants: Xavier Allamigeon, Stéphane Gaubert, Nikolas Stott.

This section presents the PhD work of Nikolas Stott. An essential part of the present work is in collaboration with Éric Goubault and Sylvie Putot (from LIX).

We develop a numerical abstract domain based on ellipsoids designed for the formal verification of switched linear systems. The novelty of this domain does not consist in the use of ellipsoids as abstractions, but rather in the fact that we overcome two key difficulties which so far have limited the use of ellipsoids in abstract interpretation. The first issue is that the ordered set of ellipsoids does not constitute a lattice. This implies that there is a priori no canonical choice of the abstraction of the union of two sets, making the analysis less predictable as it relies on the selection of good upper bounds. The second issue is that most recent works using on ellipsoids rely on LMI methods. The latter are efficient on moderate size examples but they are inherently limited by the complexity of interior point algorithms, which, in the case of matrix inequality problems, do not scale as well as for linear programming or second order cone programming problems.

We developed a new approach, in which we reduce the computation of an invariant to the determination of a fixed point, or eigenvector, of a non-linear map that provides a safe upper-approximation of the action induced by the program on the space of quadratic forms. This allows one to obtain invariants of systems of sized inaccessible by LMI methods, at the price of a limited loss of precision. A key ingredient here is the fast computation of least upper bounds in Löwner ordering, by an algebraic algorithm. This relies on the study of the geometry of the space of quadratic forms (Section 7.2.2).

The initial part of this work was described in the article [57], in which we obtained a single ellipsoidal invariant. In [16], we showed that finer disjunctive invariants, expressed as union of ellipsoids, can still be obtained by nonlinear fixed point methods in a scalable way. In [30], we developed a dual approach, which we applied to the problem of computing the joint spectral radius. We showed that an approximation of a Barabanov norm by a supremum of quadratic forms can be obtained by solving a nonlinear eigenvalue problem involving "tropical Kraus maps". The latter can be thought of as the tropical analogues of the completely positive maps appearing in quantum information. The fixed point methods in [30] allow one to solve large scale instances, unaccessible by earlier (LMI based) methods.

7.5.2. Performance evaluation of an emergency call center

Participants: Xavier Allamigeon, Vianney Boeuf, Stéphane Gaubert.

This work arose from a question raised by Régis Reboul from Préfecture de Police de Paris (PP), regarding the analysis of the projected evolution of the treatment of emergency calls (17-18-112). This work benefited from the help of LtL Stéphane Raclot, from Brigade de Sapeurs de Pompiers de Paris (BSPP), now with PP. It is part of the PhD work of Vianney Bœuf, carried out in collaboration with BSPP.

We introduced an algebraic approach which allows to analyze the performance of systems involving priorities and modeled by timed Petri nets. Our results apply to the class of Petri nets in which the places can be partitioned in two categories: the routing in certain places is subject to priority rules, whereas the routing at the other places is free choice.

We initially introduced a discrete model in [54], showing that the counter variables, which determine the number of firings of the different transitions as a function of time, are the solutions of a piecewise linear dynamical system. We showed the stationary regimes are precisely the solutions of a set of lexicographic piecewise linear equations, which constitutes a polynomial system over a tropical (min-plus) semifield of germs. However, the convergence to a stationnary regime may not occur in the discrete model. We developed in [15] a continuous time analogue of this model, involving a piecewise linear dynamical systems, and showed that it has the same stationnary regimes, avoiding some pathologies of the discrete model.

In essence, this result shows that computing stationary regimes reduces to solving tropical polynomial systems. Solving tropical polynomial systems is one of the most basic problems of tropical geometry. The latter provides insights on the nature of solutions, as well as algorithmic tools. In particular, the tropical approach allows one to determine the different congestion phases of the system. This analysis has been recovered by a probabilistic model in [42].

We applied this approach to a case study relative to the project led by Préfecture de Police de Paris, involving BSPP, of a new organization to handle emergency calls to Police (number 17), Firemen (number 18), and untyped emergency calls (number 112), in the Paris area. We combined explicit analytic computations of the different congestion phases of a simplified model and extensive simulations of a realistic an detailed model, to evaluate the performance of the center as a function of the number of operators. This analysis also suggested some ways to monitor early signs of potential congestions as well as possible correcting measures to avoid congestion.

7.5.3. Tropical models of fire propagation in urban areas

Participants: Stéphane Gaubert, Daniel Jones.

As part of the team work in the ANR project Democrite, we developed a model of fire propagation in urban areas, involving a discrete analogue of a Hamilton-Jacobi PDE. This models indicates that the fire propagates according to polyhedral ball, which is in accordance from data from historical fires (London, Chicago, or more recently Kobe).

7.5.4. Smart Data Pricing

Participants: Marianne Akian, Jean-Bernard Eytard.

This work is in collaboration with Mustapha Bouhtou (Orange Labs).

The PhD work of Jean-Bernard Eytard concerns the optimal pricing of data trafic in mobile networks. We developed a bilevel programming approach, allowing to an operator to balance the load in the network through price incentives. We showed that a subclass of bilevel programs can be solved in polynomial time, by combining methods of tropical geometry and of discrete convexity. This work is presented in [28]. In a followup work (collaboration with Gleb Koshevoy), we managed to extend these results to wider classes of bilevel problems, and to relate them to competitive equilibria problems.

7.5.5. Game theory models of decentralized mechanisms of pricing of the smart grid

Participants: Stéphane Gaubert, Paulin Jacquot.

This work is in collaboration with Nadia Oudjane and Olivier Beaude (EDF).

The PhD work of Paulin Jacquot concerns the application of game theory techniques to pricing of energy. We are developing a game theory framework for demand side management in the smart grid, in which users have movable demands (like charging an electric vehicle). We compared in particular the daily and hourly billing mechanisms. The latter, albeit more complex to analyse, has a merit in terms of incitatives, as it leads the user to move his or her consumption at off peak hours. We showed the Nash equilibrium is unique, under some assumptions, and gave theoretical bounds of the price of anarchy of the game with a hourly billing, showing this mechanism remains efficient while being more “fair” than the daily billing. We proposed and tested decentralized algorithms to compute the Nash equilibrium. These contriutions are presented In [31], [32], [44].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- Yield management methods applied to the pricing of data trafic in mobile networks. CRE (research contract) with Orange Labs (Orange Labs partner: Mustapha Bouhtou).

- Decentralized mechanisms of operation of power systems: equilibria and efficiency. A collaboration started on this topic at the fall, Nadia Oudjane, Olivier Beaude, and Riadh Zorgati from EDF-labs. This leads to the PhD work of Paulin Jacquot, supervised by Stéphane Gaubert (CIFRE PhD).

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

- Projet ANR MALTHY (Méthodes ALgébriques pour la vérification de modèles Temporisés et HYbrides), responsable T. Dang. Partenaires : Verimag, CEA LIST, Inria Rennes, Inria Saclay, VISEO/Object Direct.
- Projet ANR DEMOCRITE ("DEmonstrateur d'un MOteur de Couverture des Risques sur un TErritoire), responsable Emmanuel Lapébie (CEA). Partenaires : CEA-GRAMAT, BSPP, Inria Saclay (Maxplus), Institut PPRIME - UPR3346 (CNRS, Univ. Poitiers, ISAE-ENSMA), IPSIS, SYSTEL, ARMINES-E.M. Alès-ISR, CERDACC (Univ. de Haute-Alsace).

9.1.2. Programme Gaspard Monge pour l'Optimisation

- Projet intitulé “Méthodes tropicales pour l'optimisation”, responsable X. Allamigeon, faisant intervenir M. Akian, V. Boeuf, S. Gaubert, A.Hochart, R. Katz, et M. Skomra.

9.2. International Initiatives

9.2.1. Inria International Partners

9.2.1.1. Informal International Partners

- Collaboration with Ricardo D. Katz, CIFASIS-CONICET, Rosario (Argentina). Research invitation at CMAP during 2 months.

9.2.2. Participation in International Programs

- Collaboration with Gleb Koshevoy, Poncelet Laboratory, Moscow (research invitation of Gleb Koshevoy at CMAP during 2 months, research invitation of Stéphane Gaubert at Poncelet Laboratory during 1 week).

9.3. International Research Visitors

9.3.1. Visits of International Scientists

- Gleb Koshevoy (Russian Academy of Siences), Feb-March, 2017.
- Shmuel Friedland (University of Illinois at Chicago), one week in May 2017.
- Zheng Qu (Hong Kong University), June-July 2017
- Zheng Hua (Hong Kong University), June-July 2017
- Rajendra Bhatia (Indian Statistical Institute, New Delhi), 1 week in Dec 2017.
- Floris Claassens (University of Kent), 1 week in Dec 2017.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- Stéphane Gaubert is the coordinator of the Gaspard Monge Program for Optimization and Operations Research, a corporate sponsorhip program, operated by Fondation Mathématique Jacques Hadamard, supported by EDF, Orange and Thales, see https://www.fondation-hadamard.fr/fr/pgmo/

10.1.1.2. Member of the Organizing Committees

- S. Gaubert co-organizes the "Séminaire Parisien d'Optimisation".

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

- S. Gaubert, Chair of the PGMO days, EDF Labs Paris Saclay, Nov 13-14, 2017. https://www.fondation-hadamard.fr/fr/pgmo/pgmodays

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- S. Gaubert is member of the editorial committee of the collection Mathématiques et Applications, SMAI and Springer.
- S. Gaubert is associate editor of Linear and Multilinear Algebra.
- S. Gaubert is associate editor of RAIRO Operations research.

10.1.4. Invited Talks

- S. Gaubert
  - Noncommutative geometry: number theory, celebration of Alain Connes' 70th birthday, Shanghai, March 23 - April 7, 2017. Tropical modules, zero-sum games and non-archimedean optimization.
  - Mathematical morphology and its applications to image and signal processing Fontainebleau, Fontainebleau, May 15 – May 17, 2017. Tropical and non-linear Perron-Frobenius methods for optimal control and zero-sum games

10.1.5. Research Administration

- M. Akian :
  - Member of the "comité de liaison SMAI-MODE" since June 2015.
  - Member of the laboratory council of CMAP.
- S. Gaubert :
  - Member of the scientific council of CMAP.
- X. Allamigeon:
  - Member of the scientific committee of Inria Saclay – Ile-de-France.
  - Member of the Applied Mathematics Department committee at Ecole Polytechnique.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- M. Akian
  - Course "Markov decision processes: dynamic programming and applications" joint between (3rd year of) ENSTA and M2 "Mathématiques et Applications", U. Paris Saclay, "Optimization", and shared with Jean-Philippe Chancelier (ENPC), 15 hours each.
- X. Allamigeon

  - Petites classes et encadrement d'enseignements d'approfondissement de Recherche Opérationnelle en troisième année à l'École Polytechnique (programme d'approfondissement de Mathématiques Appliquées) (niveau M1).
  - Cours du M2 "Optimisation" de l'Université Paris Saclay, cours partagé avec Manuel Ruiz (RTE) et Dominique Quadri (LRI, Université Paris Sud).
  - Co-responsabilité du programme d'approfondissement en mathématiques appliquées (troisième année) à l'École Polytechnique.
- V. Boeuf
  - Petite classe du cours de tronc commun de 1ere année "Introduction à l'optimisation" de l'École des ponts (ENPC), niveau L3.
- J.B. Eytard
  - Cours de niveau L1-L2 à l'IUT d'Informatique d'Orsay (Univ. Paris-Sud), dans le cadre d'un monitorat (64h) (théorie des graphes, recherche opérationnelle, modélisation mathématique).
- S. Gaubert
  - Course "Systèmes à Événements Discrets", option MAREVA, ENSMP.
  - Course "Algèbre max-plus pour le contrôle optimal et les jeux" of "Parcours Optimisation, Jeux et Dynamique" (ODJ) of M2 "Mathématiques et Applications" of Paris 6 University and École Polytechnique.
  - Lecture of Operations Research, third year of École Polytechnique. The lectures notes were published as a book [63].
- M. Skomra
  - TD de mathématiques à l'UPMC.
- N. Stott
  - Cours et TD "Introduction à la programmation graphique en C++", option MAREVA et semaine Athens à l'ENSMP (École des Mines de Paris), niveau M1.

10.2.2. Supervision

- PhD in progress : Eric Fodjo, registered at École Polytechnique, since October 2013, thesis supervisor: Marianne Akian.
- PhD : Nikolas Stott, registered at École Polytechnique, since October 2014, thesis supervisor: Stéphane Gaubert, cosupervision: Xavier Allamigeon, Éric Goubault, Sylvie Putot, defended on 23 Nov 2017.
- PhD : Vianney Boeuf, registered at École Polytechnique, since October 2014, thesis supervisor: Stéphane Gaubert, cosupervision: Stéphane Raclot (BSPP), Marianne Akian, Xavier Allamigeon, defended on 18 Dec 2017.
- PhD in progress : Mateusz Skomra, registered at Univ. Paris Saclay since October 2015, thesis supervisor: Stéphane Gaubert, cosupervision: Xavier Allamigeon.
- PhD in progress : Jean-Bernard Eytard, registered at Univ. Paris Saclay since October 2015, thesis supervisor: Stéphane Gaubert, cosupervision: Marianne Akian, Mustapha Bouhtou.
- PhD in progress: Paulin Jacquot, registered at Univ. Paris Saclay since November 2016, thesis supervisor: Stéphane Gaubert, cosupervision: Nadia Oujdane, Olivier Beaude (EDF).
- PhD in progress: Benoît Tran, registered at Univ Paris-Est Marne La Vallée, since September 2017, thesis supervisor: Jean-Philippe Chancelier (ENPC), cosupervision: Marianne Akian.

10.2.3. Juries

- X. Allamigeon

  - Jury of the PhD thesis of N. Stott (Ecole Polytechnique, examiner, November 2017).
  - Jury of the PhD thesis of V. Bœuf (Ecole Polytechnique, examiner, December 2017).
- S. Gaubert
  - Member of hiring commitee (Professor position) at Paris 6 University.
  - Jury of the PhD thesis of A. Sagnier (Paris 7, examiner, 2017).
  - Jury of the PhD thesis of G. Loho (TU-Berlin, reviewer, 2017).
  - Jury of the PhD thesis of R. Hess (LAAS, 2017).
  - Jury of the PhD thesis of N. Stott (Ecole Polytechnique, examiner, November 2017).
  - Jury of the PhD thesis of V. Bœuf (Ecole Polytechnique, examiner, December 2017).

10.3. Conferences, Seminars

- M. Akian
  - Workshop "Numerical methods for optimal control problems: algorithms, analysis and applications" at INdaM, Roma, Italy, June, 19-23, 2017. Title of the talk: “Probabilistic max-plus schemes for solving Hamilton–Jacobi–Bellman equations”.
  - Atelier "Jeux dynamiques à somme nulle: temps discret, temps continu", Fréjus, from 17 to 19 oct 2017. Title of the talk: “Érgodicité des jeux à somme nulle”.
- X. Allamigeon
  - “Tropical Mathematics and its Applications” seminar, Warwick University, February 15, 2017. Title of the talk: “Tropicalization of spectrahedra”.
  - “Optimisation and Numerical Analysis” seminar, Birmingham University, February 16, 2017. Title of the talk: “Log-barrier interior-point methods are not strongly polynomial”.
  - “Computing in Tropical Geometry” workshop, Zuse Institut Berlin, May 11-12, 2017. Title of the talk: “Tropical linear optimization”.
  - 8th International Conference on Interactive Theorem Proving, Brasília, September 26-29, 2017. Title of the talk: “A Formalization of Convex Polyhedra based on the Simplex Method”.
  - “Computations and Proofs” SpecFun seminar, Saclay, October 16, 2017. Title of the talk: “Log-barrier interior-point methods are not strongly polynomial”.
- V. Boeuf
  - INFORMS APS 2017, Applied Probability Society Conference. Chicago, IL, USA, July 10-12, 2017. “An Asymptotic Analysis of Blocking in a Finite Capacity Network with Two Levels.”
- J.B. Eytard
  - Congrés annuel de la société française de Recherche Opérationnelle (ROADEF), Feb. 22-24, 2017, Metz. Title of the talk: ”Une approche tropicale de la programmation bi-niveau".
  - 15th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WIOPT) , May. 15-17, 2017, Paris. Title of the talk: "A bilevel optimization model for load balancing in mobile networks through price incentives".
  - Séminaire des doctorants du CMAP, Jun. 23, Palaiseau. Title of the talk: "A tropical approach to bilevel programming applied to a price incentives model in mobile data networks".
  - SIAM Applied Algebraic Geometry (SIAMAG), Jul. 31 - Aug. 4, Atlanta. Title of the talk: "A tropical approach to bilevel programming: application to a price incentives model in mobile data networks".

- PGMO Days, Nov. 13-14, 2016, Palaiseau. Title of the talk: “Price incentives in mobile data networks: bilevel programming, competitive equilibria and discrete convexity”.
- S. Gaubert
  - STACS, Hannover, March 2017, The operator approach to entropy games.
  - Noncommutative geometry: number theory, celebration of Alain Connes’ 70th birthday, Shanghai, March 23 - April 7, 2017. Tropical modules, zero-sum games and non-archimedean optimization.
  - Mathematical morphology and its applications to image and signal processing Fontainebleau, Fontainebleau, May 15 – May 17, 2017. Tropical and non-linear Perron-Frobenius methods for optimal control and zero-sum games
  - INDAM Workshop, Roma, June 19–23, 2017. Noncommutative aspects of dynamic programming.
  - Stony Brook Game Theory Conference, July 17-21, 2017. Nonarchimedean convexity and stochastic mean payoff games.
  - OR Berlin, Sep. 5-8, 2017. Tropical spectrahedra and stochastic mean payoff games.
  - Control and Optimization Conference On the occcasion of Frédéric Bonnans 60th birthday, EDF Labs Saclay, 15-17 Nov, 2017. Dynamic programming over noncommutative spaces
- D. Jones
  - Tropical Mathematics & its Applications, Birmingham, Nov 15 2017.“A discrete geometry model of fire propagation in urban areas”.
- M. Skomra
  - Séminaire de Géométrie Tropicale, UPMC, Paris, April 26, 2017. Title of the talk: "Tropical spectrahedra".
  - Computing in Tropical Geometry, Zuse Institute Berlin, May 11–12, 2017. Title of the talk: "Tropical spectrahedra".
  - SIAM Conference on Optimization, Vancouver, May 22–25, 2017. Title of the talk: "Solving Generic Nonarchimedean Semidefinite Programs using Stochastic Game Algorithms".
  - International Conference on Effective Methods in Algebraic Geometry (MEGA), Nice, June 12–16, 2017. Title of the talk: "The tropical analogue of the Helton–Nie conjecture is true".
  - SIAM Conference on Applied Algebraic Geometry, Atlanta, GA, Jul. 31 – Aug. 04, 2017. Title of the talk: "Tropical spectrahedra".
- C. Walsh
  - Workshop “Compactifications of buildings and symmetric spaces”, Heidelberg, May 16–17, 2017. Title of the talk: “The horofunction compactification of symmetric cones”.
  - Workshop “Order Structures, Jordan Algebras, and Geometry”, Leiden, May 29–June 2, 2017. Title of the talk: “Antitone maps and Euclidean Jordan algebras”.
  - Seminar, Ecole Polytechnique, Palaiseau, June 7, 2017. Title of the talk: “Approximability of convex bodies and volume growth in Hilbert geometries”.
  - Conference “Géométrie métrique, géométrie de Finsler”, Marseille, June 26–29, 2017. Title of the talk: “Approximability of convex bodies and volume growth in Hilbert geometries”.

11. Bibliography

Major publications by the team in recent years

[1] M. Akian, S. Gaubert, R. Bapat.*Non-archimedean valuations of eigenvalues of matrix polynomials*, in "Linear Algebra and its Applications", June 2016, vol. 498, p. 592–627, Also arXiv:1601.00438 [*DOI :* 10.1016/j.laa.2016.02.036], https://hal.inria.fr/hal-01251803.

[2] M. AKIAN, S. GAUBERT, A. GUTERMAN.*Tropical polyhedra are equivalent to mean payoff games*, in "Internat. J. Algebra Comput.", 2012, vol. 22, n^{o} 1, 1250001, 43, http://dx.doi.org/10.1142/S0218196711006674.

[3] M. AKIAN, S. GAUBERT, R. NUSSBAUM.*Uniqueness of the fixed point of nonexpansive semidifferentiable maps*, in "Transactions of the American Mathematical Society", February 2016, vol. 368, n^{o} 2, Also arXiv:1201.1536 [*DOI :* 10.1090/S0002-9947-2015-06413-7], https://hal.inria.fr/hal-00783682.

[4] M. AKIAN, S. GAUBERT, C. WALSH.*The max-plus Martin boundary*, in "Doc. Math.", 2009, vol. 14, p. 195–240.

[5] X. ALLAMIGEON, P. BENCHIMOL, S. GAUBERT.*The Tropical Shadow-Vertex Algorithm Solves Mean Payoff Games in Polynomial Time on Average*, in "ICALP 2014", Lecture Notes in Computer Science, Springer, 2014, vol. 8572, p. 89–100, http://dx.doi.org/10.1007/978-3-662-43948-7_8.

[6] X. ALLAMIGEON, P. BENCHIMOL, S. GAUBERT, M. JOSWIG.*Combinatorial simplex algorithms can solve mean payoff games*, in "SIAM J. Opt.", 2015, vol. 24, n^{o} 4, p. 2096–2117.

[7] X. ALLAMIGEON, S. GAUBERT, E. GOUBAULT, S. PUTOT, N. STOTT.*A scalable algebraic method to infer quadratic invariants of switched systems*, in "Proceedings of the International Conference on Embedded Software (EMSOFT)", 2015.

[8] J. BOLTE, S. GAUBERT, G. VIGERAL.*Definable zero-sum stochastic games*, in "Mathematics of Operations Research", 2014, vol. 40, n^{o} 1, p. 171–191, Also arXiv:1301.1967, http://dx.doi.org/10.1287/moor.2014.0666.

[9] S. GAUBERT, T. LEPOUTRE.*Discrete limit and monotonicity properties of the Floquet eigenvalue in an age structured cell division cycle model*, in "J. Math. Biol.", 2015, http://dx.doi.org/10.1007/s00285-015-0874-3.

[10] S. GAUBERT, G. VIGERAL.*A maximin characterization of the escape rate of nonexpansive mappings in metrically convex spaces*, in "Math. Proc. of Cambridge Phil. Soc.", 2012, vol. 152, p. 341–363, Also arXiv:1012.4765, http://dx.doi.org/10.1017/S0305004111000673.

[11] C. WALSH.*The horofunction boundary and isometry group of the Hilbert geometry*, in "Handbook of Hilbert Geometry", IRMA Lectures in Mathematics and Theoretical Physics, European Mathematical Society, 2014, vol. 22, https://hal.inria.fr/hal-00782827.

Publications of the year

Articles in International Peer-Reviewed Journal

[12] M. AKIAN, S. GAUBERT, A. HOCHART.*Generic uniqueness of the bias vector of finite stochastic games with perfect information*, in "Journal of Mathematical Analysis and applications", 2018, vol. 457, p. 1038-1064, https://arxiv.org/abs/1610.09651 [*DOI :* 10.1016/J.JMAA.2017.07.017], https://hal.inria.fr/hal-01425543.

[13] M. AKIAN, S. GAUBERT, A. HOCHART.*Minimax representation of nonexpansive functions and application to zero-sum recursive games*, in "Journal of Convex Analysis", February 2018, vol. 25, n^{o} 1, https://arxiv.org/abs/1605.04518 , https://hal.inria.fr/hal-01425551.

[14] M. AKIAN, S. GAUBERT, M. SHARIFY.*Log-majorization of the moduli of the eigenvalues of a matrix polynomial by tropical roots*, in "Linear Algebra and its Applications", 2017, https://arxiv.org/abs/1304.2967 - Also arXiv:1304.2967 (2013) [*DOI :* 10.1016/J.LAA.2016.11.004], https://hal.inria.fr/hal-00881196.

[15] X. ALLAMIGEON, V. BOEUF, S. GAUBERT.*Stationary solutions of discrete and continuous Petri nets with priorities*, in "Performance Evaluation", August 2017, vol. 113, p. 1 - 12, https://arxiv.org/abs/1612.07661 [*DOI :* 10.1016/J.PEVA.2017.04.007], https://hal.inria.fr/hal-01674492.

[16] X. ALLAMIGEON, S. GAUBERT, E. GOUBAULT, S. PUTOT, N. STOTT.*A Fast Method to Compute Disjunctive Quadratic Invariants of Numerical Programs*, in "ACM Transactions on Embedded Computing Systems (TECS)", October 2017, vol. 16, n^{o} 5s, p. 1-19 [*DOI :* 10.1145/3126502], https://hal.inria.fr/hal-01674495.

[17] X. ALLAMIGEON, S. GAUBERT, M. SKOMRA.*Solving generic nonarchimedean semidefinite programs using stochastic game algorithms*, in "Journal of Symbolic Computation", 2018, vol. 85, p. 25-54, An abridged version of this article appeared in the proceedings of ISSAC 2016 [*DOI :* 10.1016/J.JSC.2017.07.002], https://hal.inria.fr/hal-01674494.

[18] X. ALLAMIGEON, R. D. KATZ.*Tropicalization of facets of polytopes*, in "Linear Algebra and its Applications", June 2017, https://arxiv.org/abs/1408.6176 - Preprint arXiv:1408.6176 [*DOI :* 10.1016/J.LAA.2017.02.011], https://hal.inria.fr/hal-01096435.

[19] D. P. CASTELLA, S. GAUBERT.*Group algebra in characteristic 1 and invariant distances over finite groups*, in "Mathematische Zeitschrift", October 2017, vol. 10, p. 1-15 [*DOI :* 10.1007/S00209-017-1971-3], https://hal.inria.fr/hal-01674503.

[20] S. GAUBERT, Z. QU.*Checking the strict positivity of Kraus maps is NP-hard*, in "Information Processing Letters", February 2017, vol. 118, p. 35–43, https://arxiv.org/abs/1402.1429 - Also preprint arXiv:1402.1429 [*DOI :* 10.1016/J.IPL.2016.09.008], https://hal.inria.fr/hal-01097942.

[21] M. MACCAIG.*Exploring the complexity of the integer image problem in the max-algebra*, in "Discrete Applied Mathematics", January 2017, vol. 217, n^{o} 2, p. 261–275 [*DOI :* 10.1016/J.DAM.2016.09.016], https://hal.inria.fr/hal-01423520.

[22] A. NIV, L. ROWEN.*Dependence of tropical eigenspaces*, in "Communications in Algebra", 2017, vol. 45, n^{o} 3, p. 924-942, https://arxiv.org/abs/1504.07986 - Also arXiv:1504.07986 (2015) [*DOI :* 10.1080/00927872.2016.1172603], https://hal.inria.fr/hal-01253422.

[23] N. STOTT.*Maximal lower bounds in the Löwner order*, in "Proceedings of the AMS", 2018, https://arxiv.org/abs/1612.05664 - eprint: arXiv:1612.05664 [math.RA] [*DOI :* 10.1090/PROC/13785], https://hal.inria.fr/hal-01423497.

Invited Conferences

[24] M. AKIAN.*Probabilistic max-plus schemes for solving Hamilton-Jacobi-Bellman equations*, in "WORKSHOP INDAM: Numerical methods for optimal control problems: algorithms, analysis and applications", Roma, Italy, June 2017, https://hal.inria.fr/hal-01675069.

International Conferences with Proceedings

[25] M. AKIAN, S. GAUBERT, J. GRAND-CLÉMENT, J. GUILLAUD.*The operator approach to entropy games*, in "34th International Symposium on Theoretical Aspects of Computer Science (STACS 2017)", Hannover, Germany, March 2017, https://hal.inria.fr/hal-01676695.

[26] X. ALLAMIGEON, S. GAUBERT, M. SKOMRA.*The tropical analogue of the Helton–Nie conjecture is true*, in "MEGA 2017 - International Conference on Effective Methods in Algebraic Geometry", Nice, France, June 2017, https://hal.inria.fr/hal-01674497.

[27] X. ALLAMIGEON, R. D. KATZ.*A Formalization of Convex Polyhedra Based on the Simplex Method*, in "Interactive Theorem Proving", Brasília, Brazil, September 2017, https://arxiv.org/abs/1706.10269 , https://hal.inria.fr/hal-01673390.

[28] J. B. EYTARD, M. AKIAN, M. BOUHTOU, S. GAUBERT.*A bilevel optimization model for load balancing in mobile networks through price incentives*, in "WiOpt 2017 - 15th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks", Paris, France, IEEE, May 2017, p. 1-8 [*DOI :* 10.23919/WIOPT.2017.7959902], https://hal.inria.fr/hal-01649042.

[29] S. GAUBERT, M. JOSWIG, D. JULES.*A tropical isoperimetric inequality*, in "Proceedings of FPSAC 2017 (29th Conference on Formal Power Series and Algebraic Combinatorics, London)", London, United Kingdom, Proceedings published in Séminaire Lotharingien de Combinatoire, July 2017, vol. 78B, https://arxiv.org/abs/1611.04148 - arXiv:1611.04148, Article #27, https://hal.inria.fr/hal-01422522.

[30] S. GAUBERT, N. STOTT.*Tropical Kraus maps for optimal control of switched systems*, in "CDC 2017 - 57th IEEE Annual Conference on Decision and Control", Melbourne, Australia, December 2017, p. 1-15, https://arxiv.org/abs/1706.04471 , https://hal.inria.fr/hal-01674496.

[31] P. JACQUOT, O. BEAUDE, S. GAUBERT, N. OUDJANE.*Demand Side Management in the Smart Grid: an Efficiency and Fairness Tradeoff*, in "ISGT Europe 2017 - 7th IEEE International Conference on Innovative Smart Grid Technologies", Torino, Italy, August 2017, https://arxiv.org/abs/1711.11129 , https://hal.inria.fr/hal-01675658.

[32] P. JACQUOT, O. BEAUDE, N. OUDJANE, S. GAUBERT.*Demand Response in the Smart Grid: the Impact of Consumers Temporal Preferences*, in "IEEE International Conference on Smart Grid Communications (SmartGridComm)", Dresden, Germany, IEEE, October 2017, https://arxiv.org/abs/1711.11304 , https://hal.archives-ouvertes.fr/hal-01651263.

Conferences without Proceedings

[33] M. AKIAN, M. BOUHTOU, J. B. EYTARD, S. GAUBERT.*A tropical approach to bilevel programming: application to a price incentives model in mobile data networks*, in "SIAM Applied Algebraic Geometry (SIAMAG)", Atlanta, United States, July 2017, https://hal.inria.fr/hal-01676700.

[34] M. AKIAN, M. BOUHTOU, J. B. EYTARD, S. GAUBERT.*Une approche tropicale de la programmation bi-niveau*, in "Congrés annuel de la société française de Recherche Opérationnelle (ROADEF)", Metz, France, February 2017, https://hal.inria.fr/hal-01676696.

[35] X. ALLAMIGEON, S. GAUBERT, M. SKOMRA.*Solving Generic Nonarchimedean Semidefinite Programs Using Stochastic Game Algorithms*, in "SIAM Conference on Optimization", Vancouver, Canada, May 2017, https://hal.inria.fr/hal-01676704.

[36] X. ALLAMIGEON, S. GAUBERT, M. SKOMRA.*Tropical Spectrahedra* , in "SIAM Applied Algebraic Geometry (SIAMAG)", Atlanta, United States, July 2017, https://hal.inria.fr/hal-01676701.

[37] J. B. EYTARD, M. AKIAN, M. BOUHTOU, S. GAUBERT, G. A. KOSHEVOY.*Price incentives in mobile data networks: bilevel programming, competitive equilibria and discrete convexity*, in "PGMO Days", Palaiseau, France, November 2017, https://hal.inria.fr/hal-01676697.

Other Publications

[38] M. AKIAN, E. FODJO.*From a monotone probabilistic scheme to a probabilistic max-plus algorithm for solving Hamilton-Jacobi-Bellman equations*, July 2017, https://arxiv.org/abs/1709.09049 - Preprint, https://hal.inria.fr/hal-01675067.

[39] M. AKIAN, E. FODJO.*Probabilistic max-plus schemes for solving Hamilton-Jacobi-Bellman equations*, November 2017, working paper or preprint, https://hal.inria.fr/hal-01675068.

[40] M. AKIAN, S. GAUBERT, A. NIV.*Tropical compound matrix identities*, February 2017, https://arxiv.org/abs/1702.00980 - Preprint, https://hal.inria.fr/hal-01469638.

[41] X. ALLAMIGEON, P. BENCHIMOL, S. GAUBERT, M. JOSWIG.*Log-barrier interior point methods are not strongly polynomial*, January 2018, https://arxiv.org/abs/1708.01544 - This paper supersedes arXiv:1405.4161. 31 pages, 5 figures, 1 table, https://hal.inria.fr/hal-01674959.

[42] V. BOEUF, P. ROBERT.*A Stochastic Analysis of a Network with Two Levels of Service*, August 2017, https://arxiv.org/abs/1708.09590 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01583704.

[43] S. GAUBERT, M. MACCAIG.*Approximating the Volume of Tropical Polytopes is Difficult*, January 2018, https://arxiv.org/abs/1706.06467 - working paper or preprint, https://hal.inria.fr/hal-01675715.

[44] P. JACQUOT, O. BEAUDE, S. GAUBERT, N. OUDJANE.*Analysis and Implementation of a Hourly Billing Mechanism for Demand Response Management*, January 2018, https://arxiv.org/abs/1712.08622 - working paper or preprint, https://hal.inria.fr/hal-01675692.

[45] C. WALSH.*Order antimorphisms of finite-dimensional cones*, January 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01673777.

References in notes

[46] M. AKIAN.*Majorization inequalities for valuations of eigenvalues using tropical algebra*, in "Emerging Trends in Applied Mathematics and Mechanics (ETAMM)", Perpignan, France, May 2016, https://hal.inria.fr/hal-01425463.

[47] M. AKIAN, E. FODJO.*A probabilistic max-plus numerical method for solving stochastic control problems*, in "55th Conference on Decision and Control (CDC 2016)", Las Vegas, United States, December 2016, Also arXiv:1605.02816, https://hal.inria.fr/hal-01425344.

[48] M. AKIAN, S. GAUBERT.*Spectral theorem for convex monotone homogeneous maps, and ergodic control*, in "Nonlinear Anal.", 2003, vol. 52, n^{o} 2, p. 637–679, http://dx.doi.org/10.1016/S0362-546X(02)00170-0.

[49] M. AKIAN, S. GAUBERT, A. GUTERMAN.*Linear independence over tropical semirings and beyond*, in "Proceedings of the International Conference on Tropical and Idempotent Mathematics", G. LITVINOV, S. SERGEEV (editors), Contemporary Mathematics, American Mathematical Society, 2009, vol. 495, p. 1-38, http://www.arxiv.org/abs/0812.3496.

[50] M. AKIAN, S. GAUBERT, A. GUTERMAN.*Tropical polyhedra are equivalent to mean payoff games*, in "Internat. J. Algebra Comput.", 2012, vol. 22, n^{o} 1, 1250001, 43 [*DOI :* 10.1142/S0218196711006674], http://arxiv.org/abs/0912.2462.

[51] M. AKIAN, S. GAUBERT, A. GUTERMAN.*Tropical Cramer Determinants Revisited*, in "Tropical and Idempotent Mathematics and Applications", G. LITVINOV, S. SERGEEV (editors), Contemporary Mathematics, AMS, 2014, vol. 616, 45, See also arXiv:1309.6298, https://hal.inria.fr/hal-00881203.

[52] M. AKIAN, S. GAUBERT, A. HOCHART.*Ergodicity conditions for zero-sum games*, in "Discrete and Continuous Dynamical Systems - Series A", September 2015, vol. 35, n^{o} 9, 31, See also arXiv: 1405.4658 [*DOI :* 10.3934/DCDS.2015.35.3901], https://hal.inria.fr/hal-01096206.

[53] X. ALLAMIGEON, P. BENCHIMOL, S. GAUBERT, M. JOSWIG.*Long and winding central paths*, May 2014, Preprint arXiv:1405.4161, v2 May 2015, https://hal.inria.fr/hal-01096452.

[54] X. ALLAMIGEON, V. BOEUF, S. GAUBERT.*Performance evaluation of an emergency call center: tropical polynomial systems applied to timed Petri nets*, in "13th International Conference, Formal Modeling and Analysis of Timed Systems (FORMATS 2015)", Madrid, Spain, Formal Modeling and Analysis of Timed Systems, Springer, September 2015, vol. 9268 [*DOI :* 10.1007/978-3-319-22975-1_2], https://hal.inria.fr/hal-01248814.

[55] X. ALLAMIGEON, S. GAUBERT, É. GOUBAULT.*Inferring Min and Max Invariants Using Max-plus Polyhedra*, in "Proceedings of the 15th International Static Analysis Symposium (SAS'08)", Valencia, Spain, LNCS, Springer, 2008, vol. 5079, p. 189–204, http://dx.doi.org/10.1007/978-3-540-69166-2_13.

[56] X. ALLAMIGEON, S. GAUBERT, E. GOUBAULT.*Computing the Vertices of Tropical Polyhedra using Directed Hypergraphs*, in "Discrete Comp. Geom.", 2012, Published on line [*DOI :* 10.1007/s00454-012-9469-6], http://fr.arxiv.org/abs/0904.3436v3.

[57] X. ALLAMIGEON, S. GAUBERT, E. GOUBAULT, S. PUTOT, N. STOTT.*A scalable algebraic method to infer quadratic invariants of switched systems*, in "ACM Transactions on Embedded Computing Systems (TECS)", August 2016, vol. 15, n^{o} 4 [*DOI :* 10.1145/2932187], https://hal.inria.fr/hal-01423744.

[58] X. ALLAMIGEON, S. GAUBERT, M. SKOMRA.*Solving Generic Nonarchimedean Semidefinite Programs Using Stochastic Game Algorithms*, in "ISSAC '16: International Symposium on Symbolic and Algebraic Computation", Waterloo, France, Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation (ISSAC'16), ACM, July 2016, Also arXiv:1603.06916 [*DOI :* 10.1145/2930889.2930935], https://hal.inria.fr/hal-01422638.

[59] X. ALLAMIGEON, S. GAUBERT, M. SKOMRA.*Tropical spectrahedra*, October 2016, arXiv:1610.06746, https://hal.inria.fr/hal-01422639.

[60] P. ANDY, W. FAISAL, F. BONNANS.*MIDAS: A Mixed Integer Dynamic Approximation Scheme*, Inria, 2016, https://hal.inria.fr/hal-01401950.

[61] F. BACCELLI, G. COHEN, G.-J. OLSDER, J.-P. QUADRAT.*Synchronization and linearity: an algebra for discrete event systems*, Wiley, 1992.

[62] G. BARLES, S. MIRRAHIMI, B. PERTHAME.*Concentration in Lotka-Volterra parabolic or integral equations: a general convergence result*, in "Methods Appl. Anal.", 2009, vol. 16, n^o 3, p. 321–340, http://dx.doi.org/10.4310/MAA.2009.v16.n3.a4.

[63] F. BONNANS, S. GAUBERT.*Recherche opérationnelle. Aspects mathématiques et applications*, Ellipse, March 2016, 391, https://hal.inria.fr/hal-01422645.

[64] P. BUTKOVIČ.*Max-algebra: the linear algebra of combinatorics?*, in "Linear Algebra and its applications", 2003, vol. 367, p. 313–335.

[65] P. BUTKOVIČ.*Max-linear systems: theory and algorithms*, Springer Monographs in Mathematics, Springer-Verlag London, Ltd., London, 2010, xviii+272, http://dx.doi.org/10.1007/978-1-84996-299-5.

[66] J. COCHET-TERRASSON, G. COHEN, S. GAUBERT, M. MC GETTRICK, J.-P. QUADRAT.*Numerical computation of spectral elements in max-plus algebra*, in "Proc. of the IFAC Conference on System Structure and Control", Nantes, July 1998.

[67] G. COHEN, S. GAUBERT, J.-P. QUADRAT.*Max-plus algebra and system theory: where we are and where to go now*, in "Annual Reviews in Control", 1999, vol. 23, p. 207–219.

[68] A. CONNES, C. CONSANI.*Geometry of the arithmetic site*, 2015, arXiv:1502.05580.

[69] P. COUSOT, R. COUSOT.*Abstract Interpretation: A unified lattice model for static analysis of programs by construction of approximations of fixed points*, in "Principles of Programming Languages 4", 1977, p. 238–252.

[70] J. DE LOERA, B. STURMFELS, C. VINZANT.*The central curve in linear programming*, in "Foundations of Computational Mathematics", 2012, vol. 12, n^o 4, p. 509–540.

[71] J.-P. DEDIEU, G. MALAJOVICH, M. SHUB.*On the Curvature of the Central Path of Linear Programming Theory*, in "Foundations of Computational Mathematics", 2005, vol. 5, n^o 2, p. 145–171.

[72] A. DEZA, T. TERLAKY, Y. ZINCHENKO.*Polytopes and arrangements: diameter and curvature*, in "Operations Research Letters", 2008, vol. 36, n^o 2, p. 215–222.

[73] A. DEZA, T. TERLAKY, Y. ZINCHENKO.*Central path curvature and iteration-complexity for redundant Klee-Minty cubes*, in "Advances in applied mathematics and global optimization", New York, Adv. Mech. Math., Springer, 2009, vol. 17, p. 223–256, http://dx.doi.org/10.1007/978-0-387-75714-8_7.

[74] A. FAHIM, N. TOUZI, X. WARIN.*A probabilistic numerical method for fully nonlinear parabolic PDEs*, in "Ann. Appl. Probab.", 2011, vol. 21, n^o 4, p. 1322–1364, http://dx.doi.org/10.1214/10-AAP723.

[75] A. FATHI, A. SICONOLFI.*Existence of C^1 critical subsolutions of the Hamilton-Jacobi equation*, in "Invent. Math.", 2004, vol. 155, n^o 2, p. 363–388, http://dx.doi.org/10.1007/s00222-003-0323-6.

[76] O. FERCOQ, M. AKIAN, M. BOUHTOU, S. GAUBERT.*Ergodic control and polyhedral approaches to PageRank optimization*, in "IEEE Trans. Automat. Control", 2013, vol. 58, n^{o} 1, p. 134–148, http://dx.doi.org/10.1109/TAC.2012.2226103.

[77] W. FLEMING, W. MCENEANEY.*A max-plus based algorithm for an HJB equation of non-linear filtering*, in "SIAM J. Control and Opt.", 2000, p. 683–710.

[78] S. FOMIN, A. ZELEVINSKY.*Cluster algebras. I. Foundations*, in "J. Amer. Math. Soc.", 2002, vol. 15, n^{o} 2, p. 497–529, http://arxiv.org/abs/math.RT/0104151.

[79] S. GAUBERT, E. GOUBAULT, A. TALY, S. ZENNOU.*Static Analysis by Policy Iteration in Relational Domains*, in "Proceedings of the Proc. of the 16th European Symposium on Programming (ESOP'07)", Braga (Portugal), LNCS, Springer, 2007, vol. 4421, p. 237–252, http://dx.doi.org/10.1007/978-3-540-71316-6_17.

[80] S. GAUBERT, W. MCENEANEY, Z. QU.*Curse of dimensionality reduction in max-plus based approximation methods: theoretical estimates and improved pruning algorithms*, in "Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC 11)", Orlando, FL, USA, December 2011, p. 1054-1061, http://arxiv.org/abs/1109.5241.

[81] S. GAUBERT, A. NIV.*Tropical totally positive matrices*, December 2016, arXiv:1606.00238, https://hal.inria.fr/hal-01423747.

[82] S. GAUBERT, M. SHARIFY.*Tropical scaling of polynomial matrices*, in "Positive systems", Berlin, Lecture Notes in Control and Inform. Sci., Springer, 2009, vol. 389, p. 291–303, http://dx.doi.org/10.1007/978-3-642-02894-6_28.

[83] T. M. GAWLITZA, H. SEIDL, A. ADJÉ, S. GAUBERT, E. GOUBAULT.*Abstract interpretation meets convex optimization*, in "J. Symbolic Comput.", 2012, vol. 47, n^{o} 12, p. 1416–1446, Special issue on Invariant generation and reasoning about loops, http://dx.doi.org/10.1016/j.jsc.2011.12.048.

[84] I. M. GELFAND, M. M. KAPRANOV, A. V. ZELEVINSKY.*Discriminants, resultants and multidimensional determinants*, Modern Birkhäuser Classics, Birkhäuser Boston, Inc., Boston, MA, 2008, x+523, Reprint of the 1994 edition.

[85] S. HAMMARLING, C. J. MUNRO, F. TISSEUR.*An algorithm for the complete solution of quadratic eigenvalue problems*, in "ACM Trans. Math. Software", 2013, vol. 39, n^{o} 3, Art. 18, 19, http://dx.doi.org/10.1145/2450153.2450156.

[86] B. HEIDERGOTT, G. J. OLSDER, J. VAN DER WOUDE.*Max Plus at Work: Modeling and Analysis of Synchronized Systems: A Course on Max-Plus Algebra and Its Applications*, Princeton, 2005.

[87] A. HOCHART.*An Accretive Operator Approach to Ergodic Problems for Zero-Sum Games*, in "22nd International Symposium on Mathematical Theory of Networks and Systems (MTNS)", Minneapolis, United States, July 2016, https://hal.inria.fr/hal-01425315.

[88] A. HOCHART.*Nonlinear Perron-Frobenius theory and mean-payoff zero-sum stochastic games*, Ecole polytechnique, Palaiseau, France, November 2016, https://pastel.archives-ouvertes.fr/tel-01423953.

[89] H. ISHII, H. MITAKE.*Representation formulas for solutions of Hamilton-Jacobi equations with convex Hamiltonians*, in "Indiana Univ. Math. J.", 2007, vol. 56, n^{o} 5, p. 2159–2183, http://dx.doi.org/10.1512/iumj.2007.56.3048.

[90] I. ITENBERG, G. MIKHALKIN, E. SHUSTIN.*Tropical algebraic geometry*, Oberwolfach Seminars, Birkhäuser Verlag, Basel, 2007, vol. 35, viii+103.

[91] H. KAISE, W. M. MCENEANEY.*Idempotent expansions for continuous-time stochastic control: compact control space*, in "Proceedings of the 49th IEEE Conference on Decision and Control", Atlanta, Dec. 2010.

[92] V. KOLOKOLTSOV, V. MASLOV.*Idempotent analysis and applications*, Kluwer Acad. Publisher, 1997.

[93] B. LEMMENS, R. NUSSBAUM.*Nonlinear Perron-Frobenius theory*, Cambridge Tracts in Mathematics, Cambridge University Press, Cambridge, 2012, vol. 189, xii+323, http://dx.doi.org/10.1017/CBO9781139026079.

[94] Q. LU, M. MADSEN, M. MILATA, S. RAVN, U. FAHRENBERG, K. G. LARSEN.*Reachability Analysis for Timed Automata using Max-Plus Algebra*, in "J. Logic Alg. Prog.", 2012, vol. 81, n^{o} 3, p. 298-313.

[95] V. MASLOV.*Méthodes Operatorielles*, Edition Mir, Moscou, 1987.

[96] W. MCENEANEY, A. DESHPANDE, S. GAUBERT.*Curse-of-Complexity Attenuation in the Curse-of-Dimensionality-Free Method for HJB PDEs*, in "Proc. of the 2008 American Control Conference", Seattle, Washington, USA, June 2008.

[97] W. M. MCENEANEY, H. KAISE, S. H. HAN.*Idempotent Method for Continuous-time Stochastic Control and Complexity Attenuation*, in "Proceedings of the 18th IFAC World Congress, 2011", Milano, Italie, 2011, p. 3216-3221.

[98] W. M. MCENEANEY.*Max-plus methods for nonlinear control and estimation*, Systems & Control: Foundations & Applications, Birkhäuser Boston Inc., Boston, MA, 2006, xiv+241.

[99] W. M. MCENEANEY.*A curse-of-dimensionality-free numerical method for solution of certain HJB PDEs*, in "SIAM J. Control Optim.", 2007, vol. 46, n^{o} 4, p. 1239–1276, http://dx.doi.org/10.1137/040610830.

[100] J.-F. MERTENS, S. SORIN, S. ZAMIR.*Repeated Games*, Cambridge, 2015.

[101] G. MIKHALKIN.*Enumerative tropical algebraic geometry in $\mathbb{R}^2$*, in "J. Amer. Math. Soc.", 2005, vol. 18, n^{o} 2, p. 313–377, http://dx.doi.org/10.1090/S0894-0347-05-00477-7.

[102] R. H. MÖHRING, M. SKUTELLA, F. STORK.*Scheduling with AND/OR precedence constraints*, in "SIAM J. Comput.", 2004, vol. 33, n^{o} 2, p. 393–415, http://dx.doi.org/10.1137/S009753970037727X.

[103] A. PAPADOPOULOS.*Metric spaces, convexity and non-positive curvature*, IRMA Lectures in Mathematics and Theoretical Physics, Second, European Mathematical Society (EMS), Zürich, 2014, vol. 6, xii+309, http://dx.doi.org/10.4171/132.

[104] M. V. F. PEREIRA, L. M. V. G. PINTO.*Multi-stage stochastic optimization applied to energy planning*, in "Math. Programming", 1991, vol. 52, n^{o} 2, Ser. B, p. 359–375, http://dx.doi.org/10.1007/BF01582895.

[105] J.-E. PIN.*Tropical Semirings*, in "Idempotency", J. GUNAWARDENA (editor), Publications of the Isaac Newton Institute, Cambridge University Press, 1998, vol. 11, p. 50–69.

[106] D. REEB, M. J. KASTORYANO, M. M. WOLF.*Hilbert's projective metric in quantum information theory*, in "J. Math. Phys.", 2011, vol. 52, n$^{\mathrm{o}}$ 8, 082201, 33, http://dx.doi.org/10.1063/1.3615729.

[107] J. RICHTER-GEBERT, B. STURMFELS, T. THEOBALD.*First steps in tropical geometry*, in "Idempotent mathematics and mathematical physics", Providence, RI, Contemp. Math., Amer. Math. Soc., 2005, vol. 377, p. 289–317.

[108] G. SAGNOL, S. GAUBERT, M. BOUHTOU.*Optimal monitoring on large networks by Successive c-Optimal Designs*, in "Proceedings of the 22nd international teletraffic congress (ITC22), Amsterdam, The Netherlands, September", IEEE, 2010, http://dx.doi.org/10.1109/ITC.2010.5608717 .

[109] S. SANKARANARAYANAN, H. SIPMA, Z. MANNA.*Scalable Analysis of Linear Systems using Mathematical Programming*, in "VMCAI", LNCS, 2005, vol. 3385.

[110] C. SCHEIDERER.*Semidefinitely representable convex sets*, 2016, arXiv:1612.07048v2.

[111] R. SEPULCHRE, A. SARLETTE, P. ROUCHON.*Consensus in noncommutative spaces*, in "Proceedings of the 49th IEEE Conference on Decision and Control", Atlanta, USA, Dec 2010, p. 6596-6601, http://dx.doi.org/10.1109/CDC.2010.5717072.

[112] I. SIMON.*Limited subsets of a free monoid*, in "Proc. 19th Annual Symposium on Foundations of Computer Science", Piscataway, NJ, 1978, p. 143–150.

[113] S. SMALE.*Mathematical problems for the next century*, in "Math. Intelligencer", 1998, vol. 20, n$^{\mathrm{o}}$ 2, p. 7–15, http://dx.doi.org/10.1007/BF03025291.

[114] H. L. SMITH.*Monotone dynamical systems*, Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 1995, vol. 41, x+174, An introduction to the theory of competitive and cooperative systems.

[115] N. STOTT.*Maximal lower bounds in the Löwner order*, December 2016, eprint: arXiv:1612.05664 [math.RA], https://hal.inria.fr/hal-01423497.

[116] N. M. TRAN, J. YU.*Product-Mix Auctions and Tropical Geometry*, 2015.

[117] O. VIRO.*Dequantization of real algebraic geometry on logarithmic paper*, in "European Congress of Mathematics, Vol. I (Barcelona, 2000)", Basel, Progr. Math., Birkhäuser, 2001, vol. 201, p. 135–146, http://arxiv.org/abs/math.AG/0005163.

[118] C. WALSH.*Gauge-reversing maps on cones, and Hilbert and Thompson isometries*, 2013, 36 pages, 3 figures, http://hal.inria.fr/hal-00930929.

Project-Team XPOP

Statistical modelling for life sciences

IN COLLABORATION WITH: Centre de Mathématiques Appliquées (CMAP)

IN PARTNERSHIP WITH:
Ecole Polytechnique

RESEARCH CENTER
Saclay - Île-de-France

THEME
Modeling and Control for Life Sciences

Table of contents

Project-Team XPOP

Creation of the Team: 2016 January 01, updated into Project-Team: 2017 July 01

Keywords:

Computer Science and Digital Science:

A3.1.1. - Modeling, representation
A3.2.3. - Inference
A3.3. - Data and knowledge analysis
A3.3.1. - On-line analytical processing
A3.3.2. - Data mining
A3.3.3. - Big data analysis
A3.4.1. - Supervised learning
A3.4.2. - Unsupervised learning
A3.4.4. - Optimization and learning
A3.4.5. - Bayesian methods
A3.4.6. - Neural networks
A3.4.7. - Kernel methods
A3.4.8. - Deep learning
A5.9.2. - Estimation, modeling
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.2.2. - Numerical probability
A6.2.3. - Probabilistic methods
A6.2.4. - Statistical methods
A6.3.3. - Data processing
A6.3.5. - Uncertainty Quantification

Other Research Topics and Application Domains:

B1.1.5. - Genetics
B1.1.6. - Genomics
B1.1.9. - Bioinformatics
B1.1.11. - Systems biology
B2.2.3. - Cancer
B2.2.4. - Infectious diseases, Virology
B2.4.1. - Pharmaco kinetics and dynamics
B9.1.1. - E-learning, MOOC

1. Personnel

Research Scientist

Marc Lavielle [Team leader, Inria, Senior Researcher, HDR]

Faculty Members

Julie Josse [Ecole Polytechnique, Associate Professor]
Erwan Le Pennec [Ecole Polytechnique, Associate Professor, HDR]

Eric Moulines [Ecole Polytechnique, Professor, HDR]

Technical Staff

Yao Xu [Inria, from Nov 2017]

PhD Students

Nicolas Brosse [Ecole Polytechnique]
Wei Jiang [Ecole Polytechnique, from Oct 2017]
Mohammed Karimi [Inria]
Genevieve Robin [Ecole Polytechnique]
Marine Zulian [Dassault Systèmes]

Administrative Assistants

Hanadi Dib [Inria, from Oct 2017]
Katia Evrat [Inria, until Oct 2017]

2. Overall Objectives

2.1. Developing sound, useful and usable methods

The main objective of XPOP is to develop new sound and rigorous methods for statistical modeling in the field of biology and life sciences. These methods for modeling include statistical methods of estimation, model diagnostics and model selection as well as methods for numerical models (systems of ordinary and partial differential equations). Historically, the key area where these methods have been used is population pharmacokinetics. However, the framework is currently being extended to sophisticated numerical models in the contexts of viral dynamics, glucose-insulin processes, tumor growth, precision medicine, intracellular processes, etc.

Furthermore, an important aim of XPOP is to transfer the methods developed into software packages so that they can be used in everyday practice.

2.2. Combining numerical, statistical and stochastic components of a model

Mathematical models that characterize complex biological phenomena are defined by systems of ordinary differential equations when dealing with dynamical systems that evolve with respect to time, or by partial differential equations when there is a spatial component in the model. Also, it is sometimes useful to integrate a stochastic aspect into the dynamical system in order to model stochastic intra-individual variability.

In order to use such methods, we must deal with complex numerical difficulties, generally related to resolving the systems of differential equations. Furthermore, to be able to check the quality of a model (i.e. its descriptive and predictive performances), we require data. The statistical aspect of the model is thus critical in how it takes into account different sources of variability and uncertainty, especially when data come from several individuals and we are interested in characterizing the inter-subject variability. Here, the tools of reference are mixed-effects models.

Confronted with such complex modeling problems, one of the goals of XPOP is to show the importance of combining numerical, statistical and stochastic approaches.

2.3. Developing future standards

Linear mixed-effects models have been well-used in statistics for a long time. They are a classical approach, essentially relying on matrix calculations in Gaussian models. Whereas a solid theoretical base has been developed for such models, *nonlinear* mixed-effects models (NLMEM) have received much less attention in the statistics community, even though they have been applied to many domains of interest. It has thus been the users of these models, such as pharmacometricians, who have taken them and developed methods, without really looking to develop a clean theoretical framework or understand the mathematical properties of the methods. This is why a standard estimation method in NLMEM is to linearize the model, and few people have been interested in understanding the properties of estimators obtained in this way.

Statisticians and pharmacometricians frequently realize the need to create bridges between these two communities. We are entirely convinced that this requires the development of new standards for population modeling that can be widely accepted by these various communities. These standards include the language used for encoding a model, the approach for representing a model and the methods for using it:

- **The approach.** Our approach consists in seeing a model as hierarchical, represented by a joint probability distribution. This joint distribution can be decomposed into a product of conditional distributions, each associated with a submodel (model for observations, individual parameters, etc.). Tasks required of the modeler are thus related to these probability distributions.
- **The methods.** Many tests have shown that algorithms implemented in MONOLIX are the most reliable, all the while being extremely fast. In fact, these algorithms are precisely described and published in well known statistical journals. In particular, the SAEM algorithm, used for calculating the maximum likelihood estimation of population parameters, has shown its worth in numerous situations. Its mathematical convergence has also been proven under quite general hypotheses.
- **The language.** Mlxtran is used by MONOLIX and other modeling tools and is today by far the most advanced language for representing models. Initially developed for representing pharmacometric models, its syntax also allows it to easily code dynamical systems defined by a system of ODEs, and statistical models involving continuous, discrete and survival variables. This flexibility is a true advantage both for numerical modelers and statisticians.

3. Research Program

3.1. Scientific positioning

"Interfaces" is the defining characteristic of XPOP:

The interface between statistics, probability and numerical methods. Mathematical modelling of complex biological phenomena require to combine numerical, stochastic and statistical approaches. The CMAP is therefore the right place to be for positioning the team at the interface between several mathematical disciplines.

The interface between mathematics and the life sciences. The goal of XPOP is to bring the right answers to the right questions. These answers are mathematical tools (statistics, numerical methods, etc.), whereas the questions come from the life sciences (pharmacology, medicine, biology, etc.). This is why the point of XPOP is not to take part in mathematical projects only, but also pluridisciplinary ones.

The interface between mathematics and software development. The development of new methods is the main activity of XPOP. However, new methods are only useful if they end up being implemented in a software tool. A strong partnership with Lixoft (the spin-off company who continue developing MONOLIX) is indispensable to maintaining this positioning.

3.2. The mixed-effects models

Mixed-effects models are statistical models with both fixed effects and random effects. They are well-adapted to situations where repeated measurements are made on the same individual/statistical unit.

Consider first a single subject i of the population. Let $y_i = (y_{ij}, 1 \leq j \leq n_i)$ be the vector of observations for this subject. The model that describes the observations y_i is assumed to be a parametric probabilistic model: let $p_Y(y_i; \psi_i)$ be the probability distribution of y_i, where ψ_i is a vector of parameters.

In a population framework, the vector of parameters ψ_i is assumed to be drawn from a population distribution $p_\Psi(\psi_i; \theta)$ where θ is a vector of population parameters.

Then, the probabilistic model is the joint probability distribution

$$p(y_i, \psi_i; \theta) = p_Y(y_i|\psi_i) p_\Psi(\psi_i; \theta) \tag{17}$$

To define a model thus consists in defining precisely these two terms.

In most applications, the observed data y_i are continuous longitudinal data. We then assume the following representation for y_i:

$$y_{ij} = f(t_{ij}, \psi_i) + g(t_{ij}, \psi_i)\varepsilon_{ij} \quad , \ 1 \leq i \leq N \quad , \quad 1 \leq j \leq n_i. \tag{18}$$

Here, y_{ij} is the observation obtained from subject i at time t_{ij}. The residual errors (ε_{ij}) are assumed to be standardized random variables (mean zero and variance 1). The residual error model is represented by function g in model (2).

Function f is usually the solution to a system of ordinary differential equations (pharmacokinetic/pharmacodynamic models, etc.) or a system of partial differential equations (tumor growth, respiratory system, etc.). This component is a fundamental component of the model since it defines the prediction of the observed kinetics for a given set of parameters.

The vector of individual parameters ψ_i is usually function of a vector of population parameters $\psi_{\rm pop}$, a vector of random effects $\eta_i \sim \mathcal{N}(0, \Omega)$, a vector of individual covariates c_i (weight, age, gender, ...) and some fixed effects β.

The joint model of y and ψ depends then on a vector of parameters $\theta = (\psi_{\rm pop}, \beta, \Omega)$.

3.3. Computational Statistical Methods

Central to modern statistics is the use of probabilistic models. To relate these models to data requires the ability to calculate the probability of the observed data: the likelihood function, which is central to most statistical methods and provides a principled framework to handle uncertainty.

The emergence of computational statistics as a collection of powerful and general methodologies for carrying out likelihood-based inference made complex models with non-standard data accessible to likelihood, including hierarchical models, models with intricate latent structure, and missing data.

In particular, algorithms previously developed by POPIX for mixed effects models, and today implemented in several software tools (especially MONOLIX) are part of these methods:

- the adaptive Metropolis-Hastings algorithm allows one to sample from the conditional distribution of the individual parameters $p(\psi_i|y_i; c_i, \theta)$,
- the SAEM algorithm is used to maximize the observed likelihood $\mathcal{L}(\theta; y) = p(y; \theta)$,
- Importance Sampling Monte Carlo simulations provide an accurate estimation of the observed log-likelihood $\log(\mathcal{L}(\theta; y))$.

Computational statistics is an area which remains extremely active today. Recently, one can notice that the incentive for further improvements and innovation comes mainly from three broad directions: the high dimensional challenge, the quest for adaptive procedures that can eliminate the cumbersome process of tuning "by hand" the settings of the algorithms and the need for flexible theoretical support, arguably required by all recent developments as well as many of the traditional MCMC algorithms that are widely used in practice.

Working in these three directions is a clear objective for XPOP.

3.4. Markov Chain Monte Carlo algorithms

While these Monte Carlo algorithms have turned into standard tools over the past decade, they still face difficulties in handling less regular problems such as those involved in deriving inference for high-dimensional models. One of the main problems encountered when using MCMC in this challenging settings is that it is difficult to design a Markov chain that efficiently samples the state space of interest.

The Metropolis-adjusted Langevin algorithm (MALA) is a Markov chain Monte Carlo (MCMC) method for obtaining random samples from a probability distribution for which direct sampling is difficult. As the name suggests, MALA uses a combination of two mechanisms to generate the states of a random walk that has the target probability distribution as an invariant measure:

1. new states are proposed using Langevin dynamics, which use evaluations of the gradient of the target probability density function;
2. these proposals are accepted or rejected using the Metropolis-Hastings algorithm, which uses evaluations of the target probability density (but not its gradient).

Informally, the Langevin dynamics drives the random walk towards regions of high probability in the manner of a gradient flow, while the Metropolis-Hastings accept/reject mechanism improves the mixing and convergence properties of this random walk.

Several extensions of MALA have been proposed recently by several authors, including fMALA (fast MALA), AMALA (anisotropic MALA), MMALA (manifold MALA), position-dependent MALA (PMALA), ...

MALA and these extensions have demonstrated to represent very efficient alternative for sampling from high dimensional distributions. We therefore need to adapt these methods to general mixed effects models.

3.5. Parameter estimation

The Stochastic Approximation Expectation Maximization (SAEM) algorithm has shown to be extremely efficient for maximum likelihood estimation in incomplete data models, and particularly in mixed effects models for estimating the population parameters. However, there are several practical situations for which extensions of SAEM are still needed:

High dimensional model: a complex physiological model may have a large number of parameters (in the order of 100). Then several problems arise:

- when most of these parameters are associated with random effects, the MCMC algorithm should be able to sample, for each of the N individuals, parameters from a high dimensional distribution. Efficient MCMC methods for high dimensions are then required.
- Practical identifiability of the model is not ensured with a limited amount of data. In other words, we cannot expect to be able to properly estimate all the parameters of the model, including the fixed effects and the variance-covariance matrix of the random effects. Then, some random effects should be removed, assuming that some parameters do not vary in the population. It may also be necessary to fix the value of some parameters (using values from the literature for instance). The strategy to decide which parameters should be fixed and which random effects should be removed remains totally empirical. XPOP aims to develop a procedure that will help the modeller to take such decisions.

Large number of covariates: the covariate model aims to explain part of the inter-patient variability of some parameters. Classical methods for covariate model building are based on comparisons with respect to some criteria, usually derived from the likelihood (AIC, BIC), or some statistical test (Wald test, LRT, etc.). In other words, the modelling procedure requires two steps: first, all possible models are fitted using some estimation procedure (e.g. the SAEM algorithm) and the likelihood of each model is computed using a numerical integration procedure (e.g. Monte Carlo Importance Sampling); then, a model selection procedure chooses the "best" covariate model. Such a strategy is only possible with a reduced number of covariates, i.e., with a "small" number of models to fit and compare.

As an alternative, we are thinking about a Bayesian approach which consists of estimating simultaneously the covariate model and the parameters of the model in a single run. An (informative or uninformative) prior is defined for each model by defining a prior probability for each covariate to be included in the model. In other words, we extend the probabilistic model by introducing binary variables that indicate the presence or absence of each covariate in the model. Then, the model selection procedure consists of estimating and maximizing the conditional distribution of this sequence of binary variables. Furthermore, a probability can be associated to any of the possible covariate models.

This conditional distribution can be estimated using an MCMC procedure combined with the SAEM algorithm for estimating the population parameters of the model. In practice, such an approach can only deal with a limited number of covariates since the dimension of the probability space to explore increases exponentially with the number of covariates. Consequently, we would like to have methods able to find a small number of variables (from a large starting set) that influence certain parameters in populations of individuals. That means that, instead of estimating the conditional distribution of all the covariate models as described above, the algorithm should focus on the most likely ones.

Fixed parameters: it is quite frequent that some individual parameters of the model have no random component and are purely fixed effects. Then, the model may not belong to the exponential family anymore and the original version of SAEM cannot be used as it is. Several extensions exist:

- introduce random effects with decreasing variances for these parameters,
- introduce a prior distribution for these fixed effects,
- apply the stochastic approximation directly on the sequence of estimated parameters, instead of the sufficient statistics of the model.

None of these methods always work correctly. Furthermore, what are the pros and cons of these methods is not clear at all. Then, developing a robust methodology for such model is necessary.

Convergence toward the global maximum of the likelihood: convergence of SAEM can strongly depend on thie initial guess when the observed likelihood has several local maxima. A kind of simulated annealing version of SAEM was previously developed and implemented in MONOLIX. The method works quite well in most situations but there is no theoretical justification and choosing the settings of this algorithm (i.e. how the temperature decreases during the iterations) remains empirical. A precise analysis of the algorithm could be very useful to better understand why it "works" in practice and how to optimize it.

Convergence diagnostic: Convergence of SAEM was theoretically demonstrated under very general hypothesis. Such result is important but of little interest in practice at the time to use SAEM in a finite amount of time, i.e. in a finite number of iterations. Some qualitative and quantitative criteria should be defined in order to both optimize the settings of the algorithm, detect a poor convergence of SAEM and evaluate the quality of the results in order to avoid using them unwisely.

3.6. Model building

Defining an optimal strategy for model building is far from easy because a model is the assembled product of numerous components that need to been evaluated and perhaps improved: the structural model, residual error model, covariate model, covariance model, etc.

How to proceed so as to obtain the best possible combination of these components? There is no magic recipe but an effort will be made to provide some qualitative and quantitative criteria in order to help the modeller for building his model.

The strategy to take will mainly depend on the time we can dedicate to building the model and the time required for running it. For relatively simple models for which parameter estimation is fast, it is possible to fit many models and compare them. This can also be done if we have powerful computing facilities available (e.g., a cluster) allowing large numbers of simultaneous runs.

However, if we are working on a standard laptop or desktop computer, model building is a sequential process in which a new model is tested at each step. If the model is complex and requires significant computation time (e.g., when involving systems of ODEs), we are constrained to limit the number of models we can test in a reasonable time period. In this context, it also becomes important to carefully choose the tasks to run at each step.

3.7. Model evaluation

Diagnostic tools are recognized as an essential method for model assessment in the process of model building. Indeed, the modeler needs to confront "his" model with the experimental data before concluding that this model is able to reproduce the data and before using it for any purpose, such as prediction or simulation for instance.

The objective of a diagnostic tool is twofold: first we want to check if the assumptions made on the model are valid or not ; then, if some assumptions are rejected, we want to get some guidance on how to improve the model.

As is the usual case in statistics, it is not because this "final" model has not been rejected that it is necessarily the "true" one. All that we can say is that the experimental data does not allow us to reject it. It is merely one of perhaps many models that cannot be rejected.

Model diagnostic tools are for the most part graphical, i.e., visual; we "see" when something is not right between a chosen model and the data it is hypothesized to describe. These diagnostic plots are usually based on the empirical Bayes estimates (EBEs) of the individual parameters and EBEs of the random effects: scatterplots of individual parameters versus covariates to detect some possible relationship, scatterplots of pairs of random effects to detect some possible correlation between random effects, plot of the empirical distribution of the random effects (boxplot, histogram,...) to check if they are normally distributed, ...

The use of EBEs for diagnostic plots and statistical tests is efficient with rich data, i.e. when a significant amount of information is available in the data for recovering accurately all the individual parameters. On the contrary, tests and plots can be misleading when the estimates of the individual parameters are greatly shrunk.

We propose to develop new approaches for diagnosing mixed effects models in a general context and derive formal and unbiased statistical tests for testing separately each feature of the model.

3.8. Missing data

The ability to easily collect and gather a large amount of data from different sources can be seen as an opportunity to better understand many processes. It has already led to breakthroughs in several application areas. However, due to the wide heterogeneity of measurements and objectives, these large databases often exhibit an extraordinary high number of missing values. Hence, in addition to scientific questions, such data also present some important methodological and technical challenges for data analyst.

Missing values occur for a variety of reasons: machines that fail, survey participants who do not answer certain questions, destroyed or lost data, dead animals, damaged plants, etc. Missing values are problematic since most statistical methods can not be applied directly on a incomplete data. Many progress have been made to properly handle missing values. However, there are still many challenges that need to be addressed in the future, that are crucial for the users.

- State of arts methods often consider the case of continuous or categorical data whereas real data are very often mixed. The idea is to develop a multiple imputation method based on a specific principal component analysis (PCA) for mixed data. Indeed, PCA has been used with success to predict (impute) the missing values. A very appealing property is the ability of the method to handle very large matrices with large amount of missing entries.
- The asymptotic regime underlying modern data is not any more to consider that the sample size increases but that both number of observations and number of variables are very large. In practice first experiments showed that the coverage properties of confidence areas based on the classical

methods to estimate variance with missing values varied widely. The asymptotic method and the bootstrap do well in low-noise setting, but can fail when the noise level gets high or when the number of variables is much greater than the number of rows. On the other hand, the jackknife has good coverage properties for large noisy examples but requires a minimum number of variables to be stable enough.

- Inference with missing values is usually performed under the assumption of "Missing at Random" (MAR) values which means that the probability that a value is missing may depend on the observed data but does not depend on the missing value itself. In real data and in particular in data coming from clinical studies, both "Missing Non at Random" (MNAR) and MAR values occur. Taking into account in a proper way both types of missing values is extremely challenging but is worth investigating since the applications are extremely broad.

It is important to stress that missing data models are part of the general incomplete data models addressed by XPOP. Indeed, models with latent variables (i.e. non observed variables such as random effects in a mixed effects model), models with censored data (e.g. data below some limit of quantification) or models with dropout mechanism (e.g. when a subject in a clinical trial fails to continue in the study) can be seen as missing data models.

4. Application Domains

4.1. Precision medicine and pharmacogenomics

Pharmacogenomics involves using an individual's genome to determine whether or not a particular therapy, or dose of therapy, will be effective. Indeed, people's reaction to a given drug depends on their physiological state and environmental factors, but also to their individual genetic make-up.

Precision medicine is an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person. While some advances in precision medicine have been made, the practice is not currently in use for most diseases.

Currently, in the traditional population approach, inter-individual variability in the reaction to drugs is modeled using covariates such as weight, age, sex, ethnic origin, etc. Genetic polymorphisms susceptible to modify pharmacokinetic or pharmacodynamic parameters are much harder to include, especially as there are millions of possible polymorphisms (and thus covariates) per patient.

The challenge is to determine which genetic covariates are associated to some PKPD parameters and/or implicated in patient responses to a given drug.

Another problem encountered is the dependence of genes, as indeed, gene expression is a highly regulated process. In cases where the explanatory variables (genomic variants) are correlated, Lasso-type methods for model selection are thwarted.

There is therefore a clear need for new methods and algorithms for the estimation, validation and selection of mixed effects models adapted to the problems of genomic medicine.

A target application of this project concerns the lung cancer.

EGFR (Epidermal Growth Factor Receptor) is a cell surface protein that binds to epidermal growth factor. We know that deregulation of the downstream signaling pathway of EGFR is involved in the development of lung cancers and several gene mutations responsible for this deregulation are known.

Our objective is to identify the variants responsible for the disruption of this pathway using a modelling approach. The data that should be available for developing such model are ERK (Extracellular signal–regulated kinases) phosphorylation time series, obtained from different genetic profiles.

The model that we aim to develop will describe the relationship between the parameters of the pathway and the genomic covariates, i.e. the genetic profile. Variants related to the pathway include: variants that modify the affinity binding of ligands to receptors, variants that modify the total amount of protein, variants that affect the catalytic site,...

4.2. Oncology

In cancer, the most dreadful event is the formation of metastases that disseminate tumor cells throughout the organism. Cutaneous melanoma is a cancer, where the primary tumor can easily be removed by surgery. However, this cancer is of poor prognosis; because melanomas metastasize often and rapidly. Many melanomas arise from excessive exposure to mutagenic UV from the sun or sunbeds. As a consequence, the mutational burden of melanomas is generally high

RAC1 encodes a small GTPase that induces cell cycle progression and migration of melanoblasts during embryonic development. Patients with the recurrent P29S mutation of RAC1 have 3-fold increased odds at having regional lymph nodes invaded at the time of diagnosis. RAC1 is unlikely to be a good therapeutic target, since a potential inhibitor that would block its catalytic activity, would also lock it into the active GTP-bound state. This project thus investigates the possibility of targeting the signaling pathway downstream of RAC1.

XPOP is mainly involved in Task 1 of the project: *Identifying deregulations and mutations of the ARP2/3 pathway in melanoma patients*.

Association of over-expression or down-regulation of each marker with poor prognosis in terms of invasion of regional lymph nodes, metastases and survival, will be examined using classical univariate and multivariate analysis. We will then develop specific statistical models for survival analysis in order to associate prognosis factors to each composition of complexes. Indeed, one has to implement the further constraint that each subunit has to be contributed by one of several paralogous subunits. An original method previously developed by XPOP has already been successfully applied to WAVE complex data in breast cancer.

The developed models will be rendered user-friendly though a dedicated Rsoftware package.

This project can represent a significant step forward in precision medicine of the cutaneous melanoma.

4.3. Hemodialysis

Hemodialysis is a process for removing waste and excess water from the blood and is used primarily as an artificial replacement for lost kidney function in people with kidney failure. Side effects caused by removing too much fluid and/or removing fluid too rapidly include low blood pressure, fatigue, chest pains, leg-cramps, nausea and headaches.

Nephrologists must therefore correctly assess the hydration status in chronic hemodialysis patients and consider fluid overload effects when prescribing dialysis, according to a new study.

The fluid overload biomarker, B-type natriuretic peptide (BNP) is an important component of managing patients with kidney disease. Indeed, it is believed that each dialysis patient will have an ideal or "dry" BNP level which will accurately and reproducibly reflect their optimal fluid status.

The objective of this study is to develop a model for the BNP and the hydratation status using individual information (age, sex, ethnicity, systolic blood pressure, BMI, coronary heart disease history, ...).

The impact will be significant if the method succeeds. Indeed, it will be possible for the nephrologists to use this model for monitoring individually each treatment, in order to avoid risks of hypotension (low BNP) or overweight (high BNP).

4.4. Intracellular processes

Significant cell-to-cell heterogeneity is ubiquitously-observed in isogenic cell populations. Cells respond differently to a same stimulation. For example, accounting for such heterogeneity is essential to quantitatively understand why some bacteria survive antibiotic treatments, some cancer cells escape drug-induced suicide, stem cell do not differentiate, or some cells are not infected by pathogens.

The origins of the variability of biological processes and phenotypes are multifarious. Indeed, the observed heterogeneity of cell responses to a common stimulus can originate from differences in cell phenotypes (age, cell size, ribosome and transcription factor concentrations, etc), from spatio-temporal variations of the cell environments and from the intrinsic randomness of biochemical reactions. From systems and synthetic biology perspectives, understanding the exact contributions of these different sources of heterogeneity on the variability of cell responses is a central question.

The main ambition of this project is to propose a paradigm change in the quantitative modelling of cellular processes by shifting from mean-cell models to single-cell and population models. The main contribution of XPOP focuses on methodological developments for mixed-effects model identification in the context of growing cell populations.

- Mixed-effects models usually consider an homogeneous population of independent individuals. This assumption does not hold when the population of cells (i.e. the statistical individuals) consists of several generations of dividing cells. We then need to account for inheritance of single-cell parameters in this population. More precisely, the problem is to attribute the new state and parameter values to newborn cells given (the current estimated values for) the mother.
- The mixed-effects modelling framework corresponds to a strong assumption: differences between cells are static in time (ie, cell-specific parameters have fixed values). However, it is likely that for any given cell, ribosome levels slowly vary across time, since like any other protein, ribosomes are produced in a stochastic manner. We will therefore extend our modelling framework so as to account for the possible random fluctuations of parameter values in individual cells. Extensions based on stochastic differential equations will be investigated.
- Identifiability is a fundamental prerequisite for model identification and is also closely connected to optimal experimental design. We will derive criteria for theoretical identifiability, in which different parameter values lead to non-identical probability distributions, and for structural identifiability, which concerns the algebraic properties of the structural model, i.e. the ODE system. We will then address the problem of practical identifiability, whereby the model may be theoretically identifiable but the design of the experiment may make parameter estimation difficult and imprecise. An interesting problem is whether accounting for lineage effects can help practical identifiability of the parameters of the individuals in presence of measurement and biological noise.

4.5. Population pharmacometrics

Pharmacometrics involves the analysis and interpretation of data produced in pre-clinical and clinical trials. Population pharmacokinetics studies the variability in drug exposure for clinically safe and effective doses by focusing on identification of patient characteristics which significantly affect or are highly correlated with this variability. Disease progress modeling uses mathematical models to describe, explain, investigate and predict the changes in disease status as a function of time. A disease progress model incorporates functions describing natural disease progression and drug action.

The model based drug development (MBDD) approach establishes quantitative targets for each development step and optimizes the design of each study to meet the target. Optimizing study design requires simulations, which in turn require models. In order to arrive at a meaningful design, mechanisms need to be understood and correctly represented in the mathematical model. Furthermore, the model has to be predictive for future studies. This requirement precludes all purely empirical modeling; instead, models have to be mechanistic.

In particular, physiologically based pharmacokinetic models attempt to mathematically transcribe anatomical, physiological, physical, and chemical descriptions of phenomena involved in the ADME (Absorption - Distribution - Metabolism - Elimination) processes. A system of ordinary differential equations for the quantity of substance in each compartment involves parameters representing blood flow, pulmonary ventilation rate, organ volume, etc.

The ability to describe variability in pharmacometrics model is essential. The nonlinear mixed-effects modeling approach does this by combining the structural model component (the ODE system) with a statistical model, describing the distribution of the parameters between subjects and within subjects, as well as quantifying the unexplained or residual variability within subjects.

The objective of XPOP is to develop new methods for models defined by a very large ODE system, a large number of parameters and a large number of covariates. Contributions of XPOP in this domain are mainly methodological and there is no privileged therapeutic application at this stage.

However, it is expected that these new methods will be implemented in software tools, including MONOLIX and Rpackages for practical use.

5. Highlights of the Year

5.1. Highlights of the Year

Eric Moulines was elected at the Académie des Sciences.

The ADT *SPIX* (Analysis of very high-resolution mass spectra) was selected. This project started in November 2017 for a period of one year.

The Math-AmSud project *SaSMoTiDep* (Statistical and Stochastic modeling for time-dependent data) was selected. It begins in January 2018 for a period of two years.

6. New Software and Platforms

6.1. mlxR

KEYWORDS: Simulation - Data visualization - Clinical trial simulator
FUNCTIONAL DESCRIPTION: The models are encoded using the model coding language 'Mlxtran', automatically converted into C++ codes, compiled on the fly and linked to R using the 'Rcpp' package. That allows one to implement very easily complex ODE-based models and complex statistical models, including mixed effects models, for continuous, count, categorical, and time-to-event data.

- Contact: Marc Lavielle
- URL: http://simulx.webpopix.org/

7. New Results

7.1. Sampling from a log-concave distribution with compact support with proximal Langevin Monte Carlo

A detailed theoretical analysis of the Langevin Monte Carlo sampling algorithm was conducted when applied to log-concave probability distributions that are restricted to a convex body K. This method relies on a regularisation procedure involving the Moreau-Yosida envelope of the indicator function associated with K. Explicit convergence bounds in total variation norm and in Wasserstein distance of order 1 are established. In particular, we show that the complexity of this algorithm given a first order oracle is polynomial in the dimension of the state space.

7.2. Clustering and Model Selection via Penalized Likelihood for Different-sized Categorical Data Vectors

In this study, we consider unsupervised clustering of categorical vectors that can be of different size using mixture. We use likelihood maximization to estimate the parameters of the underlying mixture model and a penalization technique to select the number of mixture components. Regardless of the true distribution that generated the data, we show that an explicit penalty, known up to a multiplicative constant, leads to a non-asymptotic oracle inequality with the Kullback-Leibler divergence on the two sides of the inequality. This theoretical result is illustrated by a document clustering application. To this aim a novel robust expectation-maximization algorithm is proposed to estimate the mixture parameters that best represent the different topics. Slope heuristics are used to calibrate the penalty and to select a number of clusters.

7.3. Low-rank Interaction Contingency Tables

ontingency tables are collected in many scientific and engineering tasks including image processing, single-cell RNA sequencing and ecological studies. Low-rank methods have proved useful to analyze them, by facilitating visualization and interpretation. However, common methods do not take advantage of extra information which is often available, such as row and column covariates. We propose a method to denoise and visualize high-dimensional count data which directly incorporates the covariates at hand. Estimation is done by minimizing a Poisson log-likelihood and enforcing a low-rank structure on the interaction matrix with a nuclear norm penalty. We also derive theoretical upper and lower bounds on the Frobenius estimation risk. A complete methodology is proposed, including an algorithm based on the alternating direction method of multipliers, and automatic selection of the regularization parameter. The simulation study reveals that our estimator compares favorably to competitors. Then, analyzing environmental science data, we show the interpretability of the model using a biplot visualization. The method is available as an R package.

7.4. Online EM for functional data

A novel approach to perform unsupervised sequential learning for functional data is proposed. The goal is to extract reference shapes (referred to as templates) from noisy, deformed and censored realizations of curves and images. The proposed model generalizes the Bayesian dense deformable template model, a hierarchical model in which the template is the function to be estimated and the deformation is a nuisance, assumed to be random with a known prior distribution. The templates are estimated using a Monte Carlo version of the online Expectation–Maximization (EM) algorithm. The designed sequential inference framework is significantly more computationally efficient than equivalent batch learning algorithms, especially when the missing data is high-dimensional. Some numerical illustrations on curve registration problem and templates extraction from images are provided to support the methodology

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Contract with Dassault Systèmes

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

Mixed-Effects Models of Intracellular Processes: Methods, Tools and Applications (MEMIP)

Coordinator: Gregory Batt (InBio Inria team)

Other partners: InBio and IBIS Inria teams, Laboratoire Matière et Systèmes Complexes (UMR 7057; CNRS and Paris Diderot Univ.)

9.1.2. Institut National du Cancer (INCa)

Targeting Rac-dependent actin polymerization in cutaneous melanoma - Institut National du Cancer

Coordinator: Alexis Gautreau (Ecole Polytechnique)

Other partners: Laboratoire de Biochimie (Polytechnique), Institut Curie, INSERM.

9.2. International Initiatives

9.2.1. Informal International Partners

Marc Lavielle is Adjunct Professor at the Faculty of Pharmacy of Florida University.

Marc Lavielle is Adjunct Professor at the Faculty of Pharmacy of Buffalo University.

Julie Josse collaborates with Susan Holmes, Stanford University.

Eric Moulines regularly collaborates with Sean P. Meyn, University of Florida.

Geneviève Robin was recipient of a *Visiting Student Researcher Fellowship* from the France Stanford Centre for a research fellowship in the Department of Statistics at Stanford University. She worked on imputation of missing data to medical databases in a distributed framework.

9.3. International Research Visitors

9.3.1. Visits of International Scientists

Ricardo Rios, Universidad Central de Venezuela, Caracas: September 2017.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. Member of the Organizing Committees

Julie Josse and Eric Moulines were members of the Organizing Committee of the *First Data Science Summer School* at École Polytechnique (August 28th - September 1st).

Marc Lavielle was member of the Organizing Committee of the meeting *Tres dias al azar* en Cartagena, Colombia (December 14 - 16).

10.1.2. Scientific Events Selection

10.1.2.1. Member of the Conference Program Committees

Julie Josse was member of the Program Committee of the *useR!2017 meeting* in Bruxelles (June, 3-7).

10.1.3. Scientific Expertise

Marc Lavielle is member of the Scientific Committee of the High Council for Biotechnologies.

10.1.4. Research administration

Marc Lavielle is member of the Scientific Programming Committee (CPS) of the Institute Henri Poincaré (IHP).

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master : Julie Josse, Statistics with R, 48, M2, X-HEC

Master : Eric Moulines, Regression models, 36, M2, X-HEC

Engineering School : Eric Moulines, Statistics, 36, 2A, X

Engineering School : Eric Moulines, Markov Chains, 36, 3A, X

Engineering School : Erwan Le Pennec, Statistics, 36, 2A, X

Engineering School : Erwan Le Pennec, Statistical Learning, 36, 3A, X

Engineering School : Marc Lavielle, Statistics in Action, 48, 3A, X

10.2.2. Supervision

PhD in progress : Nicolas Brosse, September 2016, Eric Moulines

PhD in progress : Geneviève Robin, September 2016, Julie Josse and Eric Moulines

PhD in progress : Belhal Karimi, October 2016, Marc Lavielle and Eric Moulines

PhD in progress : Marine Zulian, October 2016, Marc Lavielle

PhD in progress : Wei Jiang , October 2017, Julie Josse and Marc Lavielle

10.3. Popularization

Marc Lavielle developed the learning platform Statistics in Action. The purpose of this online learning platform is to show how statistics (and biostatistics) may be efficiently used in practice using R. It is specifically geared towards teaching statistical modelling concepts and applications for self-study. Indeed, most of the available teaching material tends to be quite "static" while statistical modelling is very much subject to "learning by doing”.

Julie Josse participated in the jury of the "Speed data scientist" competition organized by Animath and the Société Générale. The students worked on anomaly detection: they have to identify the days of malfunctioning of the web application. They had to figure out how to anticipate breakdowns.

Julie Josse presented "How to manage missing data in R" at the meetup Rladies Paris. This meeting was associated with the Rforwards initiative, which aims to develop women’s participation in the R community.

Marc Lavielle participated to the *Jeudis de la recherche de l’X* on June 29th dedicated to “Health challenge: tools for tomorrow’s medicine”.

11. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] Y. Atchade, G. Fort, E. Moulines.*On perturbed proximal gradient algorithms*, in "Journal of Machine Learning Research", 2017, https://hal.inria.fr/hal-01668239.

[2] N. Brosse, A. Durmus, E. Moulines, M. Pereyra.*Sampling from a log-concave distribution with compact support with proximal Langevin Monte Carlo*, in "Proceedings of Machine Learning Research", 2017, vol. 65, p. 319-342, https://hal.inria.fr/hal-01648665.

[3] E. COMETS, A. LAVENU, M. LAVIELLE.*Parameter Estimation in Nonlinear Mixed Effect Models Using saemix, an R Implementation of the SAEM Algorithm*, in "Journal of Statistical Software", 2017, vol. 80, n° 3, p. 1-42 [*DOI :* 10.18637/JSS.V080.I03], https://hal.archives-ouvertes.fr/hal-01672496.

[4] R. DOUC, K. FOKIANOS, E. MOULINES.*Asymptotic properties of quasi-maximum likelihood estimators in observation-driven time series models*, in "Electronic journal of statistics ", 2017, vol. 11, n° 2, p. 2707 - 2740 [*DOI :* 10.1214/17-EJS1299], https://hal.inria.fr/hal-01668243.

[5] A. DURMUS, E. MOULINES.*Nonasymptotic convergence analysis for the unadjusted Langevin algorithm*, in "The Annals of Applied Probability : an official journal of the institute of mathematical statistics", June 2017, vol. 27, n° 3, p. 1551 - 1587 [*DOI :* 10.1214/16-AAP1238], https://hal.inria.fr/hal-01668245.

[6] M. LAVIELLE.*Pharmacometrics Models with Hidden Markovian Dynamics*, in "Journal of Pharmacokinetics and Pharmacodynamics", 2017, p. 1-15 [*DOI :* 10.1007/S10928-017-9541-1], https://hal.inria.fr/hal-01665722.

[7] F. MAIRE, E. MOULINES, S. LEFEBVRE.*Online EM for functional data*, in "Computational Statistics and Data Analysis", July 2017, vol. 111, p. 27 - 47 [*DOI :* 10.1016/J.CSDA.2017.01.006], https://hal.inria.fr/hal-01668241.

[8] N. M. NGUYEN, S. LE CORFF, E. MOULINES.*Particle rejuvenation of Rao-Blackwellized sequential Monte Carlo smoothers for conditionally linear and Gaussian models*, in "EURASIP Journal on Advances in Signal Processing", December 2017, vol. 2017:54, p. 1-15 [*DOI :* 10.1186/S13634-017-0489-5], https://hal.inria.fr/hal-01668374.

[9] H.-T. WAI, J. LAFOND, A. SCAGLIONE, E. MOULINES.*Decentralized Frank–Wolfe Algorithm for Convex and Nonconvex Problems*, in "IEEE Transactions on Automatic Control", November 2017, vol. 62, n° 11, p. 5522 - 5537 [*DOI :* 10.1109/TAC.2017.2685559], https://hal.inria.fr/hal-01668247.

International Conferences with Proceedings

[10] H.-T. WAI, J. LAFOND, A. SCAGLIONE, E. MOULINES.*Fast and privacy preserving distributed low-rank regression*, in "2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)", New Orleans, United States, 2017, https://hal.inria.fr/hal-01668252.

Other Publications

[11] N. BROSSE, A. DURMUS, E. MOULINES.*Normalizing constants of log-concave densities*, November 2017, working paper or preprint, https://hal.inria.fr/hal-01648666.

[12] N. BROSSE, A. DURMUS, E. MOULINES, S. SABANIS.*The Tamed Unadjusted Langevin Algorithm*, November 2017, working paper or preprint, https://hal.inria.fr/hal-01648667.

[13] E. DERMAN, E. LE PENNEC.*Clustering and Model Selection via Penalized Likelihood for Different-sized Categorical Data Vectors*, September 2017, https://arxiv.org/abs/1709.02294 - working paper or preprint, https://hal.inria.fr/hal-01583692.

[14] E. GAUTIER, E. LE PENNEC.*Adaptive estimation in the nonparametric random coefficients binary choice model by needlet thresholding*, November 2017, https://arxiv.org/abs/1106.3503 - working paper or preprint, https://hal.inria.fr/inria-00601274.

[15] G. ROBIN, J. JOSSE, E. MOULINES, S. SARDY.*Low-rank Interaction Contingency Tables*, September 2017, https://arxiv.org/abs/1703.02296 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01482773.