Activity Report 2013

Section New Results

5. New Results

5.1. RNA

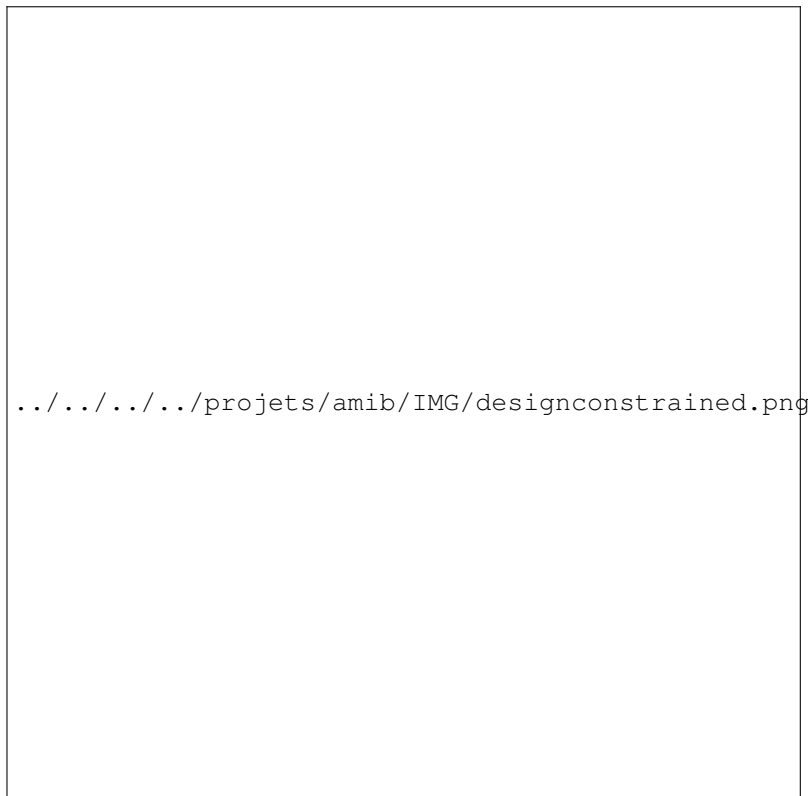../../../../projets/amib/IMG/designconstrained.png

Figure 3. Language-theoretical constructs for the constrained design of RNAs. Starting from a secondary structure, the language of sequences compatible with base-pairing constraints is modeled as a context-free grammar (Left), while forced and forbidden motifs (here, {AA} is forbidden, and {AGC, GG} are forced) can be modeled by a dedicated automaton (Right).

5.1.1. RNA design through random generation

Extensive experiments revealed a drift of existing software towards sequences with a high G+C-content. Relying on our random generation methods, we showed how to control this distributional bias in sequences using a multidimensional Boltzmann sampling [30], [22]. We also explored the combination of random generation (global sampling) and local search into a novel category of *glocal* approaches, yielding promising results.

Finally, we explored language-theoretic constructs, namely products of finite-state automata and context-free languages, to force or forbid the presence of identified functional motifs within designed sequences [33].

5.1.2. *Towards 3D modeling of large molecules*

Ab initio research benefited from our works on research and classification of RNA structural motifs [63]. Significant progress towards the *ab initio* prediction of the 3D structure of large RNAs were achieved. This problem is beyond the scope of current approaches and we proposed a promising coarse-grained approach based on game theory [13] that scales up to several hundreds of bases.

5.1.3. *Fast-fourier transform for riboswitch*

In the field of RNA computational biology, many algorithms use dynamic programming to partition the folding landscape according to a set of structural parameters. More precisely, the goal is to compute the number (resp. cumulated Boltzmann weight) $c_{p_1,p_2,p_3...}$ of secondary structures having p_i occurrences of some structural parameter P_i, where P_i may denote the distance to a reference structure, the number of # helices, base-pairs...The resulting algorithms, although polynomial in theory, are usually unusable in practice, particularly due to their unreasonable complexities (typically $\Theta(n^{3+2k})/\Theta(n^{2+k})$ time/memory for k parameters) and the intrinsic difficulties one encounters while trying to distribute their computation over multiple processors (highly connected dependency graph).

In collaboration with P. Clote's group (Boston College), we have described generic algorithmic principles to dramatically decrease these complexities, and make this class of algorithms practical. The main idea is to capture the partitioned space within a large polynomial, which can typically be efficiently evaluated (typically in $\Theta(n^3)$) as soon as the parameters are additive. One can then perform (possibly in parallel) $\Theta(n^k)$ independent evaluations of the polynomial, and use the Discrete Fourier Transform to recover the coefficients in $\Theta(k \cdot n^k \cdot \log{(n)})$ time. Applying these principles to the `RNAbor` algorithm, whose complexities were in $\Theta(n^5)/\Theta(n^3)$, we obtained an novel $\Theta(n^4)/\Theta(n^2)$ (parallelizable in $\Theta(n^3)/\Theta(n^2)$ time/memory on $m \to \infty$ processors), we obtained a novel algorithm to detect bistable thermodynamic structures, such as riboswitches, which we presented at `Recomb`'13 [32].

5.2. Sequences

5.2.1. *Random generation*

The random generation of decomposable combinatorial structures, pioneered by P. Flajolet in the 80s, provides an elegant, yet powerful, framework to model and sample the objects which appear in computational biology. Random samples can then be used to assert the significance of a given observable when closed form formulae are difficult to obtain.

Messenger RNAs (mRNAs) encode proteins, but may also independently feature structured motifs which are crucial to recoding and alternative splicing mechanisms. In order to predict such motifs, the stability of smaller regions within a given mRNA must be compared to that of sequences generated with respect to a **background model** which, at the same time, preserves the encoded amino-acid sequence and the capacity of the overall sequence to form a stable fold (proxy-ed by the dinucleotide composition). Using multidimensional Boltzmann sampling, we have revisited the underlying – well-defined, yet never solved exactly – random generation problem, and provided the first unbiased and practical algorithm for the problem [27]. The algorithm, developed in collaboration with McGill and Université de Montréal (Canada), has linear time complexity as soon as a small tolerance (typically $\Theta(1/\sqrt{n})$) on the composition is allowed.

Some other biological objects, such as RNA secondary structures, naturally appear with probabilities which are poorly modeled by the uniform distribution. To better model such objects, Denise *et al* [3] have introduced the **weighted distribution**, and adapted classic random generation algorithms such that each object within a given combinatorial family can be generated with respect to it. However, the exponentially increasing probability ratio between the most and least probable object sometimes leads to a large degree of redundancy within generated sets . To work around this issue, and generate non-redundant sets of objects, we have proposed a sequential algorithm with deterministically avoids any previously generated word, without introducing any bias in the generation [17].

../../../../projets/amib/IMG/sparcs.png

Figure 4. Workflow of our `NASP` *pipeline [27]: An assessment of significantly (un)-structured regions in protein-coding RNAs can be achieved through a dinucleotide-preserving random generation of sequences encoding the same protein.*

../../../../projets/amib/IMG/nonredgen.png

Figure 5. A uniform random generation of words avoiding a predefined set of words can be achieved using a dedicated data structure, leading to a careful correction of the emission probabilities. Enriching the set of forbidden words after each generation, one obtains a non-redundant generation algorithm [17].

Besides, in collaboration with the Fortesse group at LRI, we developed a new divide and conquer algorithm for the random generation of words of regular languages, and we performed a complete benchmarking of all state-of-the-art methods dedicates to this problem [56].

```
../../../../projets/amib/IMG/RNAPyro.png
```

Figure 6. While simultaneously sequencing the genome of a (microbial) community, Next-Generation Sequencing techniques produce small genomic fragments, whose diversity arises from a combination of genetic variants and sequencing errors. We used knowledge of the RNA secondary structure to develop a pre-filter that detects and corrects post-mapping sequencing errors.

5.2.2. Next Generation Sequencing (NGS)

As a side-product of our previous collaborative studies with J. Waldispühl (McGill, Canada), focusing on sequence/structure relationship in RNA, we revisited the problem of detecting and correcting RNA sequences obtained using pyrosequencing techniques. Indeed, ribosomal RNAs are often used to estimate the population diversity within a microbiome, and sequencing errors may lead to biased estimates. In this context, we

investigated whether a complete knowledge of the RNA secondary structure could be exploited to detect and correct errors in NGS reads.

To that end, we introduced a probabilistic model, defined over all sequences at maximal distance d from the input read and their respective folding. This model captures both the stability of the induced fold and its compatibility with a reference multiple sequence alignment. We designed a linear-time inside/outside algorithm to compute exactly the probability that a given position is mutated in the ensemble. Our initial implementation, presented at RECOMB'13 [29] and published an extended version in *Journal of Computational Biology* [23], revealed encouraging results, and we plan to combine it with a population diversity estimator to test its potential in a metagenomics context.

5.2.3. *Combinatorics of motifs*

An algorithm for pvalue computation has been proposed in [44] that takes into account a Hiddden Markov Model and an implementation, SUFPREF, has been realized (http://server2.lpm.org.ru/bio).

Combinatorics of clumps have been extensively studied, leading to the definition of the so-called *canonic clumps*. It is shown in [28] that they contain the necessary information needed to calculate, approximate, and study probabilities of occurrences and asymptotics. This motivates the development of a *clump automaton*. It allows for a derivation of pvalues, decreasing the space and time complexity of the generating function approach or previous weighted automata.

Large deviations approximations are needed for very rare events, e.g. very small pvalues, as Gaussian approximations are known not to be applicable. In [21], combinatorial properties of words allow to provide an explicit and tractable formula for the tail distribution with a low space and time complexity and a guaranteed tightness. Double strands counting problem is addressed where dependencies between a sequence and its complement plays a fundamental role. A large deviation result is also provided for a set of small sequences, with non-identical distributions. Possible applications are the search of cis-acting elements in regulatory sequences that may be known, for example from ChIP-chip or ChipSeq experiments, as being under a similar regulatory control. In a recent internship at LIX, F. Pirot detected a Chi-like motif in Archae genome.

In a collaboration with AlFarabi University, where M. Régnier acts as a foreign co-advisor), word statistics were used to identify mRNA targets for miRNAs involved in various cancers [8], [9].

5.3. 3D Modelling and Interactions

Transmembrane beta-barrel proteins (TMB) account for 20 to 30% of identified proteins in a genome but, due to difficulties with standard experimental techniques, they are only 2% of the RCSB Protein Data Bank. Therefore, we study and design algorithmic solutions addressing the secondary structure, an abstraction of the 3D conformation of a molecule, that only retains the contacts between its residues. Although this representation may disregard some of the fine details of the molecule conformation, it still retains the general architecture of molecules, and is especially useful in the study of RiboNucleic Acids (RNAs) and transmembrane beta-barrel proteins (TMB). The latter class of proteins accounts for 20 to 30% of identified proteins in a genome but, due to difficulties with standard experimental techniques, they constitute only 2As TMB perform many vital functions, the prediction of their structure is a challenge for life sciences, while the small number of known structures prohibits knowledge-based methods for structure prediction. As TMBs are strongly structured objects, model based methodologies [26], [25] are an interesting alternative to these conventional methods. The efficiently obtained 3D structures provide a good model for further 3D and interaction analyses.

In a recent work [34], we focused on the identification of protein-protein complexes based on the putative interaction between pairs of proteins as the sole source of information. From the results obtained on *E. coli*, we started working on the prediction of multi-body protein complexes from sequence information alone.

In our protein-RNA project, we managed to obtain the first learning results. We optimized the RosettaDock scores and showed that such an optimization cannot be done efficiently without expert knowledge. The first results are to be presented at EGC in 2014 [61].

5.3.1. Large scale cross-docking study of the specificity of protein-protein interactions

The year 2013 saw the conclusion of a long-term collaboration, involving A. Carbone (UPMC) and A. Lopes (IGM, Paris XI). In a recent paper published in the prestigious *Plos Computational Biology* [16] journal, we showed that combining coarse-grain molecular cross-docking simulations and binding site predictions based on evolutionary sequence analysis is a viable route to identify true interacting partners for hundreds of proteins with a variate set of protein structures and interfaces. Also, we realized a large-scale analysis of protein binding promiscuity and provided a numerical characterization of partner competition and level of interaction strength for about 28000 false-partner interactions. Finally, we demonstrated that binding site prediction is useful to discriminate native partners, but also to scale up the approach to thousands of protein interactions. This study was based on a large computational effort made by thousands of internet users helping the World Community Grid over a period of 7 months.

5.4. Data Integration

Work performed in the Data Integration axis this year has been dedicated to the design and implementation of a new approach to reduce the complexity of scientific workflow structures. More precisely, we focused on the presence of "anti-patterns" in the workflow structures, idiomatic structures that lead to over-complicated design. We have then proposed the *DistilFlow* method and a tool for automatically detecting such anti-patterns and replacing them with different patterns which result in a reduction in the workflow's overall structural complexity [10] (BMC Journal paper accepted, published early 2014). This work has been performed in close collaboration with the Taverna group from the University of Manchester.

DistilFlow is part of J. Chen's thesis who has defended his PhD on October 11th, 2013 [7] and is now back to China as a research assistant in Lanzhou University.

5.5. Systems Biology

Systems Biology includes the study of interaction networks such as gene regulatory, metabolic, or signaling networks. It involves both designing the topology of the networks and predicting their dynamic and spatiotemporal aspects. It requires the import of concepts from across various disciplines and crosstalk between theory, benchwork, modelling and simulation.

5.5.1. Topological analysis of metabolic networks

In [73] we have developed a biclustering algorithm for elementary flux modes that is based on the Agglomeration of Common Motifs (ACoM). This allows a drastic diminution of the number of less significant fluxes and a kind of factorization of most important fluxes, yielding an algorithm running in quadratic time in the number of elementary flux modes.

We applied this algorithm to describe the decomposition into elementary flux modes of the central carbon metabolism in *Bacillus subtilis* and of the yeast mitochondrial energy metabolism. For *Bacillus subtilis*, a specific inhibition on the second domain of the lipoamide dehydrogenase (pdhD) component of pyruvate dehydrogenase complex that leads to the loss of all fluxes was exhibited [20]. Such a conclusion is not predictable in the classical approach.

5.5.2. Evolution of metabolic networks

A collaboration with IGM on the evolution of metabolic networks is ungoing. We aim at understanding how such networks would emerge over time among the variety of species, and how these changes could be responsible for characteristic life traits. Our methodology to characterize the evolutionary origin of the enzymatic repertoire of different fungal groups relies on machine learning. Preliminary results were presented at JOBIM 2013 [35].

5.5.3. *Signaling networks*

Our goal is to help the understanding of signaling pathways involving (GPCR) and to provide means to semi-automatically construct the signaling networks. Our method takes into account various kinds of biological experiments and their origin and automatically builds and draws the inferred network. Comparing the automatically deduced network with an already known fragment of the FSHR network allowed us to obtain new interesting hypotheses that are currently experimentally tested by biologists, our collaborators from INRA-BIOSin Tours. In the next months, experimental data for some GPCR (FSH, 5HT2 et 5HT4) will be prepared by BIOS and IGF (Montpellier), in the context of a GPCRNET ANR project.

Besides, in collaboration with K. Inoue, through the NII International Internship Program, we have studied the System Biology Graphical Notation language, a standard for expressing molecular networks, especially signaling networks, and proposed a translation of SBGN-AF into a logical formalism [31].

5.5.3.1. Modelling with Hsim

In a collaboration of P. Amar with microbiologists, the group of Marie-Joëlle Virolle from the *Institut de Génétique et de Microbiologie*, a first explicative model was proposed for the sigmoidicity of the shape of the survival curve of bacteria (*S. lividans*) having a antibiotic resistance gene, expressed at different levels, in presence of a constant concentration of antibiotics [24], [6], [18], [41].

This is particularly important since this method of inclusion of an antibiotics resistance gene to report the activity of its promoter is widely used in the streptomyces community.

5.5.3.2. Cancer and metabolism

It is shown in M. Behzadi's PhD thesis that most systems have very stable behaviours and that even large variations of their chemical characteristics do not affect the nature of the equilibria. This very general situation has been discovered by simulation but in some cases it is even possible to prove it mathematically.

Our collaborators M. Israël and L. Schwartz have listed more than a hundred tentative such bifurcations that we intend to study systematically. A preliminary study of the mitotic cycle with L. Paulevé has also put in evidence the strong influence of the pH of the cell on its capacity to duplicate. The PhD thesis of E. Bigan, co-directed by S. Daoudi (Univ. Denis Diderot) and J.-M. Steyaert investigates the generic properties of such complex systems and confirms that the ones we have already studied are not exceptions [43]. Some prospective cases are studied in [14].

<p align="center" style="color:red">AVIZ Project-Team</p>

6. New Results

6.1. Hybrid-Image Visualizations

Participants: Petra Isenberg [correspondant], Pierre Dragicevic, Wesley Willett, Anastasia Bezerianos, Jean-Daniel Fekete.

We investigated hybrid-image visualization for data analysis in large-scale viewing environments. Hybrid-image visualizations blend two different visual representations into a single static view, such that each representation can be perceived at a different viewing distance. Our work was motivated by data analysis scenarios that incorporate one or more displays with sufficiently large size and resolution to be comfortably viewed by different people from various distances. Hybrid-image visualizations can be used, in particular, to enhance overview tasks from a distance and detail-in-context tasks when standing close to the display. By taking advantage of humans' perceptual capabilities, hybrid-image visualizations do not require tracking of viewers in front of a display. Moreover, because hybrid-images use a perception-based blending approach, visualizations intended for different distances can each utilize the entire display. In our paper we contributed a design space, discussed the perceptual rationale for our work, provided examples and a set of techniques for hybrid-image visualizations, and described tools for designing hybrid-image visualizations. We will also release software that will help in the construction of hybrid-image visualizations.

```
../../../../projets/aviz/IMG/hybrid.jpg
```

Figure 8. Overview of a treemap showing a subset of the tree of life with a hybrid image visualization. Larger structures are clearly visible from far but do not interfere with reading detail when up close (see in-set).

6.2. Visualization for Interactive Displays

Participants: Tobias Isenberg [correspondant], Petra Isenberg.

Because the access to and analysis of information is becoming increasingly important anywhere and at any time, researchers have begun to investigate the role of interactive displays as data analysis platforms. Visualization applications play a crucial role in data analysis and development of dedicated systems and tools for small to large interactive displays to support such application contexts is underway. We contribute a systematic and quantitative assessment of the literature from ten different venues, an open repository of papers, and a code-set that can be used to categorize the research space [22]. We found just over 100 publications at the intersection of interactive surfaces and visualization in our careful examination of 10 different publication venues related to the topic. We found that research has so far largely focused on the development of interaction techniques, for multi-touch tabletop devices, and 2D spatial and abstract visualizations. Together, all publications addressed a wide spectrum of research questions and, given the many possible combinations of interactive surfaces and visualization, the research space is still wide open. While several projects developed applications for data analysis with visualization on interactive surfaces, their availability in practice is still rare. Commercial companies and open-source communities have begun to provide ported versions of their products/tools for tablets and mobile phones (e.g., Tableau Mobile 4 and KiwiViewer 5), showing the need for visualization application on surfaces. Nevertheless, the support for data analysis tasks on these and other interactive surfaces can certainly still be improved—a lot more research with respect to the development and evaluation of the fundamentals of data exploration and analysis is needed for interactive displays. Some example directions of future work in this context were outlined in research agendas published in the journal IEEE Computer [24] as well as in IEEE Computer Graphics and Applications [21].

In a specific project investigated an interaction design concept for exploratory 3D data visualization that marries direct-touch interaction with stereoscopic vision. The design is inspired by the mental mapping that occurs for mouse interaction where the physical control space is mapped through a mental rotation to the display space. Similarly, we explore touch interaction on a monoscopic tablet, mapped through a mental rotation to the stereoscopic display space. Because this mental mapping can become increasingly skewed we show when and how to re-synchronize the views (see Figure 9).

../../../../projets/aviz/IMG/david-tablet.jpg

Figure 9. Tablet-based navigation of a stereoscopically displayed 3D dataset.

6.3. Visualization for Soccer Analysis

Participants: Charles Perin, Romain Vuillemot, Jean-Daniel Fekete [correspondant].

A new generation of soccer data is now available, as some companies (http://www.optasports.com/) collect and provide extensive data covering almost all professional soccer championships, with a wealth of multivariate information related to time, player positions, and types of action, to name a few. Currently, most analysis on such data relate to statistics on individual players or teams. For instance, statistics on "team ball possession" and "number of goal attempts for team A or B" are popular on websites, TV and newspapers and often accompanied by bar charts or plots on a soccer field. However, soccer analysts we collaborated with consider that quantitative analysis alone does not convey the right picture of the game, as context, player positions and phases of player actions are the most relevant aspects.

SoccerStories [] (Figure 10) is a visualization interface to support analysts in exploring soccer data and communicating interesting insights that we designed to support the current practice of soccer analysts and to enrich it, both in the analysis and communication stages. Our system provides an overview+detail interface of game phases, and their aggregation into a series of connected visualizations, each visualization being tailored for actions such as a series of passes or a goal attempt. To evaluate our tool, we ran two qualitative user studies on recent games using SoccerStories with data from one of the world's leading live sports data providers. The first study resulted in a series of four articles on soccer tactics, by a tactics analyst, who said he would not have been able to write these otherwise. The second study consisted in an exploratory follow-up to investigate design alternatives for embedding soccer phases into word-sized graphics. For both experiments, we received a very enthusiastic feedback and participants consider further use of SoccerStories to enhance their current workflow. This article received a Best Paper Honorable Mention in VIS 2013.

We also explored how spectators of a live soccer game can collect detailed data while watching the game [46]. Our motivation arouse from the lack of free detailed sport data, contrasting with the large amount of simple statistics collected for every popular games and available on the web. Assuming many spectators carry a smart phone during a game, we implemented a series of input interfaces for collecting data in real time. In a user study, we asked participants to use those interfaces to perform tracking tasks such as locating players in the field, qualifying ball passes, and naming the player with ball while watching a video clip of a real soccer game. Our two main results are 1) the crowd can collect detailed and fairly complex data in real-time with reasonable quality while each participant is assigned a simple task, and 2) a set of design implications for crowd-powered interfaces to collect live sport data. We also discuss the use of such data into SoccerStories, and the design implications coming with the visual communication of missing and uncertain detailed data.

Finally, we presented R2S2 [45] in the SportVis workshop (VIS 2013), a hybrid visualization technique as an intermediate step between Rank Chart and Slope Graph to better understand and analyze team evolutions during soccer championships. Currently used rank tables for soccer are relative (ranked-based) and do not convey the absolute difference between teams. R2S2 provides a way to visualize these differences using the Slope Graph technique (value-based). By interactively setting the parameters of R2S2, we make the distance between teams appear, minimizing the overlaps caused by the Slope Graph technique.

More information about these projects is available at http://www.aviz.fr/soccer.

6.4. Interaction Model for Visualizations Beyond the Desktop

Participants: Yvonne Jansen [correspondant], Pierre Dragicevic.

We introduced an interaction model for beyond-desktop visualizations that combines the visualization reference model with the instrumental interaction paradigm. Beyond-desktop visualizations involve a wide range of emerging technologies such as wall-sized displays, 3D and shape-changing displays, touch and tangible input, and physical information visualizations. While these technologies allow for new forms of interaction, they are often studied in isolation. New conceptual models are needed to build a coherent picture of what has been done and what is possible. We described a modified pipeline model where raw data is processed into a visualization and then rendered into the physical world. Users can explore or change data by directly manipulating visualizations or through the use of instruments. Interactions can also take place in the physical world outside the visualization system, such as when using locomotion to inspect a large scale visualization. Through case

../../../../projets/aviz/IMG/soccer.png

Figure 10. Using SoccerStories: (a) navigating among soccer phases of a game; (b) mapping a phase on a focus soccer field; (c) exploring the phase by grouping actions into tailored visualizations; and (d) communicating using Sportlines embed into text.

studies we illustrated how this model can be used to describe both conventional and unconventional interactive visualization systems, and compare different design alternatives.

6.5. Network Visualization

Participants: Benjamin Bach [correspondant], Basak Alper, Andre Spritzer, Emmanuel Pietriga, Nathalie Henry-Riche, Tobias Isenberg, Jean-Daniel Fekete.

Although much research has been done on finding efficient ways to visualize different kinds of networks (social networks, computer networks, brain networks, etc), many problems are still open. Rather than trying to find optimal layouts, we focus on novel representation and navigation techniques to explore such networks. Our research focusses on three major problems: (i) heterogeneous networks, (ii) comparison of graphs, (iii) dynamic networks, and (iv) generating networks for controlled user evaluations.

Heterogeneous Networks: Heterogeneous networks are networks with multiple node and edge types, such as ontologies in the Semantic Web. Ontologies also provide means to describe type hierarchies on node and edge types as well as other set theoretic relations between these types. Such heterogeneous networks can be explored on two levels, the model (types and possible relations between types), and instances (actual nodes and their connections). In order to allow for interactive exploration of such data, we extended the NodeTrix technique [67] to visualize ontologies. Our prototype is called *OntoTrix* [11] (Figure 11 (a), which allows for reorganizing matrices by splitting and merging them, traversing node and edge hierarchies and visualize different types of connection between nodes.

Graph Comparision: Analyzing brain networks, which can represent anatomical fibers as well as functional correlation between brain regions, is complex in many ways. By analysing brain scientists tasks we concluded that many user tasks can actually performed by comparing two networks. In [28](Figure 11 (b)) we design and discuss several ways to compare two weighted graphs and finally compare the two most promising designs in a controlled user study. We found that our encoding for adjacency matrices outperforms the one for node-link diagrams, even for sparse networks. The implications for brain analysis tools are manyfold and our results generalize to other domains that are concerned with comparing (dense and weighted) networks.

Dynamic Networks: A very common technique to explore dynamic networks are animations and small multiples, each of which being supporting different tasks, while falling short on others. With *GraphDiaries* [10](Figure 11 (c)), we design an interface based on the combination of both techniques while offering flexible temporal navigation techniques as well as enhanced perceptive feedback to understand changes between time steps. *GraphDiaries* supports further navigation techniques such as temporal aggregation, direct difference views and layout adjustment. While *GraphDiaries* is highly extensible, its techniques are designed to be integrated in existing visualization tools.

While animations and the techniques in *GraphDiaries* are useful for many networks, dense dynamic networks are still an important open problem. We hence generalized the idea of matrices to visualize temporal networks, by describing a visualization and interaction model based on the space time cube metaphor (Figure 11 (d)). In analogy with the physical world, this *Matrix Cube* can be manipulated and decomposed in order to explore the network, while the cube model serves as a consistent visual and mental model of the data and visualization. We implemented an interface called *Cubix* that allows us to perform simple view switches and decomposition operations in the cube. Cubix and the Matrix Cube was evaluated with two experts, an astronomer and brain scientist, exploring their own real world data. With the Matrix Cube and its decomposition operations, we are able to visualize and navigate within very dense dynamic networks such as brain networks, trading flows and technical networks. The design space of possible visualizations that the Matrix Cube and its operations offer is both, huge but structured at the same time. It allows us to explore many future designs.

As part of the effort of visualizing publications and work of Jean-Daniel Fekete, we designed a visualization to show his collaborations over the past years, relating his papers and his collaborators [29](Figure 11 (e)). A poster was presented at part of a poster submission to IEEE Vis, 2013 in Atlanta. To the best of our knowledge, no such technique to visualize any sort of dynamic ego networks have been published so far.
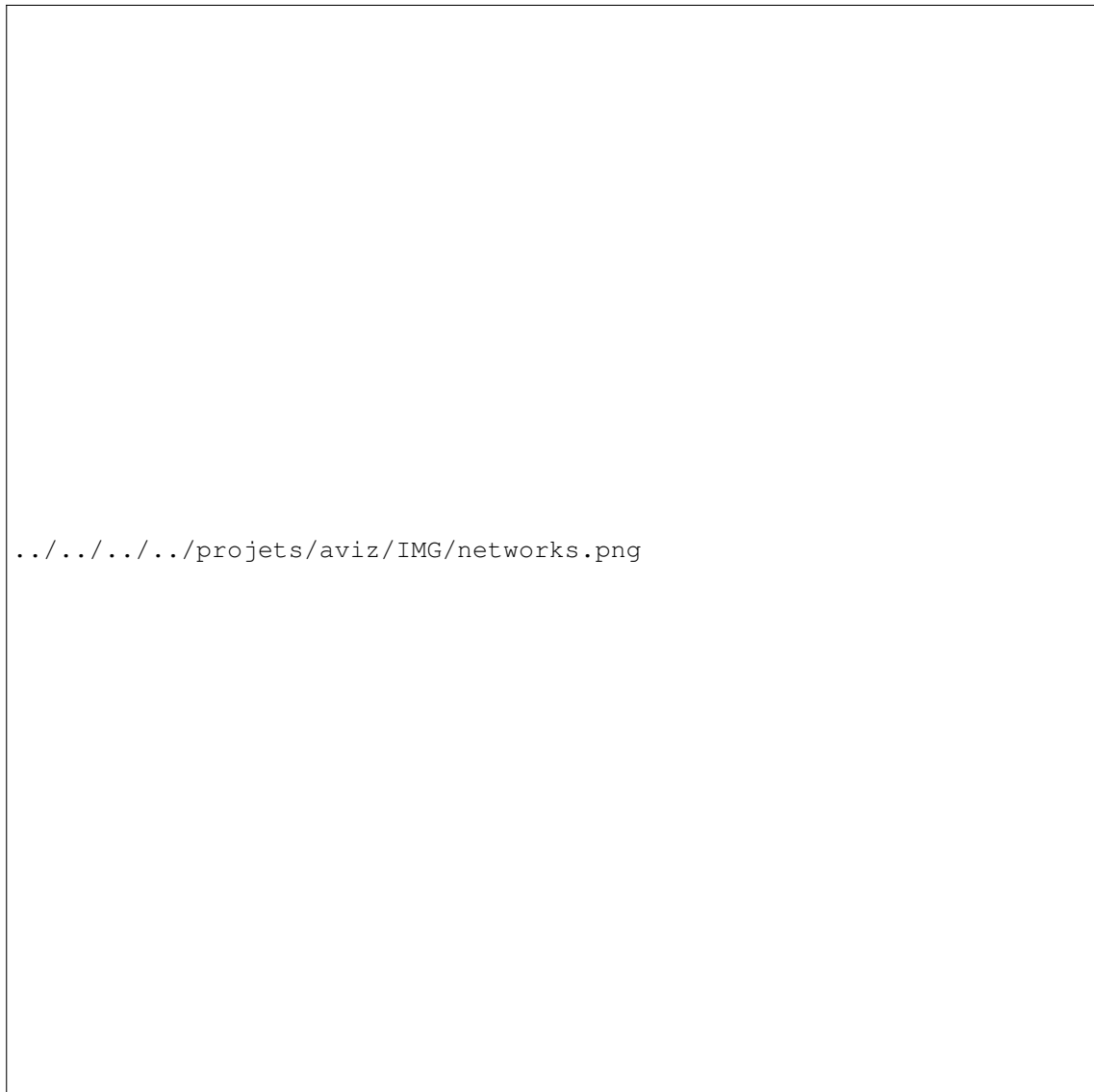
../../../../projets/aviz/IMG/networks.png

Figure 11. Examples for network visualizations

Network Generation: As any controlled user study, evaluating network visualizations requires control over the data. However, it is hard to found real world data with the desired properties and in reasonable amount for a controlled user study. Synthetic data can help but the output of random graph generators is hard to control and hardly resembles actual real-world data. With GraphCuisine [56] we present and interactive approach to generate graphs. In an iterative process, the computer generates suggestions while the user selects her preferred graphs and graph measures.

6.6. GridVis: Visualisation of Island-Based Parallel Genetic Algorithms

Participants: Waldo Cancino [correspondant], Hugo Gilbert, Benjamin Bach, Evelyne Lutton, Pierre Collet.

Island Model parallel genetic algorithms rely on various migration models and their associated parameter setting. A fine understanding of how the islands interact and exchange informations is an important issue for the design of efficient algorithms. GridVis, is an interactive tool that has been developed for visualising the exchange of individuals and the propagation of fitness values between islands. GridVis has been developed in Java, to monitor how the islands communicate: when and how much individuals of which fitness they effectively exchange during a run. We model the computer cluster that is running the island model, as dynamic network and use an adjacency matrix to show the relations (exchange between individuals) between nodes (computers) in the cluster (Figure 12 (a)). Several experiments have been performed on a grid and on a cluster to evaluate GridVis' ability to visualise the activity of each machine and the communication flow between machines. Experiments have been made on the optimisation of a Weierstrass function using the EASEA language, with two schemes: a scheme based on uniform islands and another based on specialised islands (Exploitation, Exploration and Storage Islands).

../../../../projets/aviz/IMG/figure1_labelled.png

Figure 12. Visualisation of a grid with 20 machines: Each computer in the cluster appears twice in the matrix, once as row and once as column. Cells inside the matrix indicate information about the interaction of computers during evolution, for example, the amount of individuals exchanged (read from row to column). Similar to heat maps activity (exchange of individuals) is mapped to darkness (dark cells indicate high exchange, bright cells show low exchange).

<p style="text-align:center">**COMETE Project-Team**</p>

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. Differential privacy with general metrics.

Differential privacy can be interpreted as a bound on the distinguishability of two generic databases, which is determined by their Hamming distance: the distance in the graph determined by the adjacency relation (two databases are adjacent if they differ for one individual).

In [21] we lifted the restriction relative to the Hamming graphs and we explored the implications of differential privacy when the indistinguishability requirement depends on an arbitrary notion of distance. We showed that we can express, in this way, (protection against) kinds of privacy threats that cannot be naturally represented with the standard notion. We gave an intuitive characterization of these threats in terms of Bayesian adversaries, which generalizes the characterization of (standard) differential privacy from the literature. Next, we revisited the well-known result on the non-existence of universally optimal mechanisms for any query other than counting queries. We showed that in our setting, for certain kinds of distances, there are many more queries for which universally optimal mechanisms exist: Notably sum, average, and percentile queries. Finally, we showed some applications in various domains: statistical databases where the units of protection are groups (rather than individuals), geolocation, and smart metering.

6.1.2. Privacy for location-based services.

The growing popularity of location-based services, allowing unknown/untrusted servers to easily collect and process huge amounts of users' information regarding their location, has recently started raising serious concerns about the privacy of this kind of sensitive information. In [19] we studied geo-indistinguishability, a formal notion of privacy for location-based services that protects the exact location of a user, while still allowing approximate information - typically needed to obtain a certain desired service - to be released.

Our privacy definition formalizes the intuitive notion of protecting the user's location within a radius r with a level of privacy that depends on r. We presented three equivalent characterizations of this notion, one of which corresponds to a generalized version [21] of the well-known concept of differential privacy. Furthermore, we presented a perturbation technique for achieving geo-indistinguishability by adding controlled random noise to the user's location, drawn from a planar Laplace distribution. We demonstrated the applicability of our technique through two case studies: First, we showed how to enhance applications for location-based services with privacy guarantees by implementing our technique on the client side of the application. Second, we showed how to apply our technique to sanitize location-based sensible information collected by the US Census Bureau.

6.1.3. Relation between differential privacy and quantitative information flow.

Differential privacy is a notion that has emerged in the community of statistical databases, as a response to the problem of protecting the privacy of the database's participants when performing statistical queries. The idea is that a randomized query satisfies differential privacy if the likelihood of obtaining a certain answer for a database x is not too different from the likelihood of obtaining the same answer on adjacent databases, i.e. databases which differ from x for only one individual.

In [13], we analyzed critically the notion of differential privacy in light of the conceptual framework provided by the Rényi min information theory. We proved that there is a close relation between differential privacy and leakage, due to the graph symmetries induced by the adjacency relation. Furthermore, we considered the utility of the randomized answer, which measures its expected degree of accuracy. We focused on certain kinds of utility functions called "binary", which have a close correspondence with the Rényi min mutual information. Again, it turns out that there can be a tight correspondence between differential privacy and utility, depending on the symmetries induced by the adjacency relation and by the query. Depending on these symmetries we can also build an optimal-utility randomization mechanism while preserving the required level of differential privacy. Our main contribution was a study of the kind of structures that can be induced by the adjacency relation and the query, and how to use them to derive bounds on the leakage and achieve the optimal utility.

6.1.4. A differentially private mechanism of optimal utility for a region of priors

Differential privacy (already introduced in the previous sections) is usually achieved by using mechanisms that add random noise to the query answer. Thus, privacy is obtained at the cost of reducing the accuracy, and therefore the utility, of the answer. Since the utility depends on the user's side information, commonly modeled as a prior distribution, a natural goal is to design mechanisms that are optimal for every prior. However, it has been shown in the literature that such mechanisms do not exist for any query other than counting queries.

Given the above negative result, in [22] we considered the problem of identifying a restricted class of priors for which an optimal mechanism does exist. Given an arbitrary query and a privacy parameter, we geometrically characterized a special region of priors as a convex polytope in the priors space. We then derived upper bounds for utility as well as for min-entropy leakage for the priors in this region. Finally we defined what we call the tight-constraints mechanism and we discussed the conditions for its existence. This mechanism has the property of reaching the bounds for all the priors of the region, and thus it is optimal on the whole region.

6.1.5. Compositional analysis of information hiding

Systems concerned with information hiding often use randomization to obfuscate the link between the observables and the information to be protected. The degree of protection provided by a system can be expressed in terms of the probability of error associated to the inference of the secret information. In [14] we considered a probabilistic process calculus to specify such systems, and we studied how the operators affect the probability of error. In particular, we characterized constructs that have the property of not decreasing the degree of protection, and that can therefore be considered safe in the modular construction of these systems. As a case study, we applied these techniques to the Dining Cryptographers, and we derived a generalization of Chaum's strong anonymity result.

In [26], a similar framework was proposed for reasoning about the degree of differential privacy provided by such systems. In particular, we investigated the preservation of the degree of privacy under composition via the various operators. We illustrated our idea by proving an anonymity-preservation property for a variant of the Crowds protocol for which the standard analyses from the literature are inapplicable. Finally, we made some preliminary steps towards automatically computing the degree of privacy of a system in a compositional way.

6.1.6. Preserving differential privacy under finite-precision semantics

The approximation introduced by finite-precision representation of continuous data can induce arbitrarily large information leaks even when the computation using exact semantics is secure. Such leakage can thus undermine design efforts aimed at protecting sensitive information. For instance, the standard approach to achieve differential privacy (introduced in previous sections) is the addition of noise to the true (private) value. To date, this approach has been proved correct only in the ideal case in which computations are made using an idealized, infinite-precision semantics. In [23], we analyzed the situation at the implementation level, where the semantics is necessarily finite-precision, i.e. the representation of real numbers and the operations on them are rounded according to some level of precision. We showed that in general there are violations of the differential privacy property, and we studied the conditions under which we can still guarantee a limited (but, arguably, totally acceptable) variant of the property, under only a minor degradation of the privacy

level. Finally, we illustrated our results on two cases of noise-generating distributions: the standard Laplacian mechanism commonly used in differential privacy, and a bivariate version of the Laplacian recently introduced in the setting of privacy-aware geolocation.

6.1.7. *Metrics for differential privacy in concurrent systems*

Many protocols for protecting confidential information have involved randomized mechanisms and a nondeterministic behavior (such as the Dining Cryptographers protocol or the Crowds protocol). In [28], we investigate techniques for proving differential privacy in the context of concurrent systems which contain both probabilistic and nondeterministic behaviors. Our motivation stems from the work of Tschantz et al., who proposed a verification method based on proving the existence of a stratified family of bijections between states, that can track the privacy leakage, ensuring that it does not exceed a given leakage budget. We improve this technique by investigating state properties which are more permissive and still imply differential privacy. We consider three pseudometrics on probabilistic automata: The first one is essentially a reformulation of the notion proposed by Tschantz et al. The second one is a more liberal variant, still based on the existence of a family of bijections, but relaxing the relation between them by integrating the notion of amortization, which results into a more parsimonious use of the privacy budget. The third one aims at relaxing the bijection requirement, and is inspired by the Kantorovich-based bisimulation metric proposed by Desharnais et al. We cannot adopt the latter notion directly because it does not imply differential privacy. Thus we propose a multiplicative variant of it, and prove that it is still an extension of weak bisimulation. We show that for all the pseudometrics the level of differential privacy is continuous on the distance between the starting states, which makes them suitable for verification. Moreover we formally compare these three pseudometrics, proving that the latter two metrics are indeed more permissive than the first one, but incomparable with each other, thus constituting two alternative techniques for the verification of differential privacy.

6.1.8. *Unlinkability*

Unlinkability is a privacy property of crucial importance for several systems (such as RFID or voting systems). Informally, unlinkability states that, given two events/items in a system, an attacker is not able to infer whether they are related to each other. However, in the literature we find several definitions for this notion, which are apparently unrelated and shows a potentially problematic lack of agreement. In [20] we shed new light on unlinkability by comparing different ways of defining it and showing that in many practical situations the various definitions coincide. It does so by (a) expressing in a unifying framework four definitions of unlinkability from the literature (b) demonstrating how these definitions are different yet related to each other and to their dual notion of "inseparability" and (c) by identifying conditions under which all these definitions become equivalent. We argued that the conditions are reasonable to expect in identification systems, and we prove that they hold for a generic class of protocols.

6.1.9. *Trust in anonymity networks*

Trust metrics are used in anonymity networks to support and enhance reliability in the absence of verifiable identities, and a variety of security attacks currently focus on degrading a user's trustworthiness in the eyes of the other users. In [16] we have presented an enhancement of the Crowds anonymity protocol via a notion of trust which allows crowd members to route their traffic according to their perceived degree of trustworthiness of each other member of the crowd. Such trust relations express a measure of an individual's belief that another user may become compromised by an attacker, either by a direct attempt to corrupt or by a denial-of-service attack. Our protocol variation has the potential of improving the overall trustworthiness of data exchanges in anonymity networks, which cannot normally be taken for granted in a context where users are actively trying to conceal their identities. Using such formalization, in the paper we have then analyzed quantitatively the privacy properties of the protocol under standard and adaptive attacks.

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. Models and Emerging Trends of Concurrent Constraint Programming

The *Concurrent constraint programming (ccp)* paradigm focuses on information access and therefore it is suited for this new era of concurrent systems. Ccp singles out the fundamental aspects of asynchronous systems whose agents (or processes) evolve by accessing information in a global medium, represented as constraints over the variables of the system. Agents communicate by posting and querying partial information in the medium. This covers a vast variety of systems as those arising in biological phenomena, reactive systems, net- centric computing and the advent of social networks and cloud computing. In [17] we surveyed the main applications, developments and current trends of ccp.

6.2.2. Efficient computation of program equivalence for confluent concurrent constraint programming

The development of algorithms and automatic verification procedures for ccp have hitherto been far too little considered. To the best of our knowledge there is only one existing verification algorithm for the standard notion of ccp program (observational) equivalence. In [25] we first showed that this verification algorithm has an exponential-time complexity even for programs from a representative sub-language of ccp; the summation-free fragment (ccp+). We then significantly improved on the complexity of this algorithm by providing two alternative polynomial-time decision procedures for ccp+ program equivalence. Each of these two procedures has an advantage over the other. One has a better time complexity. The other can be easily adapted for the full language of ccp to produce significant state space reductions. The relevance of both procedures derives from the importance of ccp+. This fragment, which has been the subject of many theoretical studies, has strong ties to first-order logic and an elegant denotational semantics, and it can be used to model real-world situations. Its most distinctive feature is that of confluence, a property we exploit to obtain our polynomial procedures.

6.2.3. Abstract Interpretation of Temporal Concurrent Constraint Programs

Timed concurrent constraint programming (tcc) is a declarative model for concurrency offering a logic for specifying reactive systems, i.e. systems that continuously interact with the environment. The universal tcc formalism (utcc) is an extension of tcc with the ability to express mobility. Here mobility is understood as communication of private names as typically done for mobile systems and security protocols. In [15] we considered the denotational semantics for tcc, and we extended it to a "collecting" semantics for utcc based on closure operators over sequences of constraints. Relying on this semantics, we formalized a general framework for data flow analyses of tcc and utcc programs by abstract interpretation techniques. The concrete and abstract semantics we proposed are compositional, thus allowing us to reduce the complexity of data flow analyses. We showed that our method is sound and parametric with respect to the abstract domain. Thus, different analyses can be performed by instantiating the framework. We illustrated how it is possible to reuse abstract domains previously defined for logic programming to perform, for instance, a groundness analysis for tcc programs. We showed the applicability of this analysis in the context of reactive systems. Furthermore, we made also use of the abstract semantics to exhibit a secrecy flaw in a security protocol. We also showed how it is possible to make an analysis which may show that tcc programs are suspension free. This can be useful for several purposes, such as for optimizing compilation or for debugging.

6.2.4. Foundations of Probabilistic Concurrent Systems

In [24] we introduced a formal proof system for compositional verification of probabilistic concurrent processes. Properties are expressed using a probabilistic modal μ-calculus, and the proof system is formulated as a sequent calculus in which sequents are given a quantitative interpretation. A key feature is that the probabilistic scenario is handled by introducing the notion of Markov proof, by which each proof in the

system is interpreted as a Markov Decision Process, with the proof only considered valid in the case that the value of the MDP is zero.

5. New Results

5.1. Optimality conditions in Pontryagin form for optimal control problems

Participants: Joseph Frédéric Bonnans, Xavier Dupuis, Laurent Pfeiffer.

5.1.1. Necessary conditions

In the paper [31], we state and prove first- and second-order necessary conditions in Pontryagin form for optimal control problems with pure state and mixed control-state constraints. We say that a Lagrange multiplier of an optimal control problem is a Pontryagin multiplier if it is such that Pontryagin's minimum principle holds, and we call optimality conditions in Pontryagin form those which only involve Pontryagin multipliers. Our conditions rely on a technique of partial relaxation, and apply to Pontryagin local minima.

5.1.2. Sufficient conditions

In the paper [32], we consider sufficient conditions. More precisely, given a reference feasible trajectory of an optimal control problem, we say that the quadratic growth property for bounded strong solutions holds if the cost function of the problem has a quadratic growth over the set of feasible trajectories with a bounded control and with a state variable sufficiently close to the reference state variable. Our sufficient second-order optimality conditions in Pontryagin form ensure this property and ensure *a fortiori* that the reference trajectory is a bounded strong solution. Our proof relies on a decomposition principle, which is a particular second-order expansion of the Lagrangian of the problem.

5.1.3. Shooting Approach to Optimal Control Problems

Participant: Joseph Frédéric Bonnans.

In the paper [24] we give an overview of the shooting technique for solving deterministic optimal control problems. This approach allows to reduce locally these problems to a finite dimensional equation. We first recall the basic idea, in the case of unconstrained or control constrained problems, and show the link with second-order optimality conditions and the analysis or discretization errors. Then we focus on two cases that are now better understood: state constrained problems, and affine control systems. We end by discussing extensions to the optimal control of a parabolic equation.

5.2. Applications of deterministic optimal control problems

5.2.1. Optimization of running strategies based on anaerobic energy and variations of velocity

Participant: Joseph Frédéric Bonnans.

In the report [29] we present new models, numerical simulations and rigorous analysis for the optimization of the velocity in a race. In a seminal paper, Keller [74], [75] explained how a runner should determine his speed in order to run a given distance in the shortest time. We extend this analysis, based on the equation of motion and aerobic energy, to include a balance of anaerobic energy (or accumulated oxygen deficit) and an energy recreation term when the speed decreases. We also take into account that when the anaerobic energy gets too low, the oxygen uptake cannot be maintained to its maximal value. Using optimal control theory, we obtain a proof of Keller's optimal race, and relate the problem to a relaxed formulation, where the propulsive force represents a probability distribution rather than a value function of time. Our analysis leads us to introduce a bound on the variations of the propulsive force to obtain a more realistic model which displays oscillations of the velocity. Our numerical simulations qualitatively reproduce quite well physiological measurements on real runners. We show how, by optimizing over a period, we recover these oscillations of speed. We point out that our numerical simulations provide in particular the exact instantaneous anaerobic energy used in the exercise.

5.2.2. *Optimal control of leukemic cell population dynamics*
Participant: Xavier Dupuis.

In the paper [33] we discuss the optimal co-administration of two drugs for some acute myeloid leukemias (AML), and we are looking for in vitro protocols as a first step. This issue can be formulated as an optimal control problem. The dynamics of leukemic cell populations in culture is given by age-structured partial differential equations, which can be reduced to a system of delay differential equations, and where the controls represent the action of the drugs. The objective function relies on eigenelements of the uncontrolled model and on general relative entropy, with the idea to maximize the efficiency of the protocols. The constraints take into account the toxicity of the drugs. We present in this paper the modeling aspects, as well as theoretical and numerical results on the optimal control problem that we get.

5.2.3. *Contrast imaging problem in nuclear magnetic resonance*
Participant: Pierre Martinon.

In collaboration with team McTAO (Sophia), we studied in [25] and [36] the contrast imaging problem in nuclear magnetic resonance, modeled as Mayer problem in optimal control. The optimal solution can be found as an extremal, solution of the Maximum Principle and analyzed with the techniques of geometric control. A first synthesis of locally optimal solutions is given in the single-input case, with some preliminary results in the bi-input case. We conducted a comprehensive numerical investigation of the problem, using a combination of indirect shooting (HAMPATH software) and direct method (BOCOP), with a moment-based (LMI) technique to estimate the global optimum.

../../../../projets/commands/IMG/bi-blood-1p1.png

Figure 2. Contrast in quantum control for NMR - Oxygenated / deoxygenated blood

5.2.4. *Optimizing the anaerobic digestion of microalgae in a coupled process*
Participant: Pierre Martinon.

In collaboration with the Inra-Inria team MODEMIC (Montpellier), we studied in [30] a bio-reactor system describing the coupling of a culture of micro-algae and an anaerobic digester. Our aim is to optimize the production of methane in the digester during a certain number of days with respect to the dilution rate (the input flow of micro-algae in the digester). The mathematical model for the dynamics of the two reactors takes into account a periodic day-night model of the light in the culture of micro-algae, and a chemostat model for the digester. We first prove existence and attraction of periodic solutions for a one day period, and we apply Pontryagin's Maximum Principle (PMP) in order to characterize optimal controls. We provide numerical simulations for different light models, by a direct method that we refine using an indirect shooting. We also investigate the dependence of the optimal cost with respect to the ratio of the volumes of the two tanks. Finally, we investigate the optimal strategies over a large number of days without periodic constraints, and compared the mean cost to the optimal cost over one period.

Figure 3. Coupled bio-reactor for micro-algae digestion - Attraction property

5.2.5. *Design of optimal experiments for parameter estimation of microalgae growth models*
Participant: Pierre Martinon.

In collaboration with team BIOCORE (Sophia), we investigated in [27] techniques of Optimal Experiment Design for microalgae growth models. In order to have microalgae growth models that are useful for prediction and process optimization, reliable parameters need to be provided. This reliability implies a careful design of experiments that can be exploited for parameter estimation. OED techniques can provide guidelines for the design of experiments with high informative content that allow an accurate parameter estimation. We study a real experimental device devoted to evaluate the effect of temperature and light on microalgae growth. On the basis of a mathematical model of the experimental system, the optimal experiment design problem was solved

as an optimal control problem. E-optimal experiments were obtained by using two discretization approaches, namely sequential and simultaneous. The results showed that an adequate parameterization of the experimental inputs provided optimal solutions very close to those provided by the simultaneous discretization. Simulation results showed the relevance of determining optimal experimental inputs for achieving an accurate parameter estimation.

Figure 4. Experimental apparatus for the study of micro-algae growth (Ifremer)

5.2.6. *Controllability and optimal strokes for N-link microswimmer*
Participant: Pierre Martinon.

In [39] we focus on the N-link swimmer, a generalization of the classical Purcell swimmer. We use the simplification of the Resistive Force Theory to derive the motion equation for the swimmer in a fluid with a low Reynolds number. We prove that the swimmer is controllable in the whole plane when it is composed by more than 3 sticks and for almost every set of stick lengths. As a direct result, we show that there exists an optimal swimming strategy which leads to minimize the time to reach a desired configuration. Numerical experiments on the case of the Purcell swimmer suggest that the optimal strategy is periodic, i.e. composed of a sequence of identical strokes. Our results indicate that this candidate for an optimal stroke indeed gives a better speed than the classical Purcell stroke. Future directions for this work include the design of robotic micro-swimmers, as well as investigation of the movement of swimming micro-organisms.

5.3. Hamilton-Jacobi (HJ) approach

5.3.1. *Dynamic programming and error estimates for stochastic control with Max cost*
Participants: Olivier Bokanowski, Athena Picarelli, Hasnaa Zidani.

../../../../projets/commands/IMG/swimmer_xf.png

Figure 5. Purcell (3-link) swimmer - Purcell vs optimal stroke

The paper [35] is concerned with stochastic optimal control for a running maximum cost. A direct approach based on dynamic programming techniques is studied leading to the characterization of the value function as the unique viscosity solution of a second order Hamilton-Jacobi-Bellman (HJB) equation with an oblique derivative boundary condition. A general numerical scheme is proposed and a convergence result is provided. Error estimates are obtained for the semi-Lagrangian scheme. These results can apply to the case of lookback options in finance. Moreover, optimal control problems with maximum cost arise in the characterization of the reachable sets for a system of controlled stochastic differential equations. Some numerical simulations on examples of reachable analysis are included to illustrate our approach.

5.3.2. *Optimal feedback control of undamped wave equations by solving a HJB equation*
Participant: Hasnaa Zidani.

An optimal finite-time horizon feedback control problem for (semi linear) wave equations is studied in [42]. The feedback law can be derived from the dynamic programming principle and requires to solve the evolutionary Hamilton-Jacobi-Bellman (HJB) equation. Classical discretization methods based on finite elements lead to approximated problems governed by ODEs in high dimensional space which makes infeasible the numerical resolution by HJB approach. In the present paper, an approximation based on spectral elements is used to discretize the wave equation. The effect of noise is considered and numerical simulations are presented to show the relevance of the approach

5.3.3. *Transmission conditions on interfaces for Hamilton-Jacobi-Bellman equations*
Participants: Hasnaa Zidani, Zhiping Rao.

The works [43], [28] deal with deterministic control problems where the dynamic and the running cost can be completely different in two (or more) complementary domains of the space IR^N. As a consequence, the dynamics and running cost present discontinuities at the interfaces of these domains. This leads to a complex interplay that has to be analyzed among transmission conditions to "glue" the propagation of the value function on the interfaces. Several questions arise: how to define properly the value function(s) and what is (are) the right Bellman Equation(s) associated with this problem?. In the case of a simple geometry (namely when the space IR^N is partitioned into two subdomains separated with an interface which is assumed to be a regular hypersurface without any connectedness requirement), [43] discuss different conditions on the hyperplane where the dynamic and the running cost are discontinuous, and the uniqueness properties of the Bellman problem are studied. In this paper it is used a dynamical approach, namely instead of working with test functions, the accent is put on invariance properties of an augmented dynamics related to the integrated control system. The comparison principle is accordingly based, rather than on (semi)continuity of the Hamiltonian appearing in the Hamilton–Jacobi–Bellman equation, on some weak separation properties of this dynamics with respect to the stratification. A more general situation where the space is partionned on several domains is also analyzed in [28].

5.3.4. *Singular perturbation of optimal control problems on multi-domains*
Participants: Nicolas Forcadel, Hasnaa Zidani.

The goal of the paper[38] is to study a singular perturbation problem in the framework of optimal control on multi-domains. We consider an optimal control problem in which the controlled system contains a fast and a slow variables. This problem is reformulated as an Hamilton-Jacobi-Bellman (HJB) equation. The main difficulty comes from the fact that the fast variable lives in a multi-domain. The geometric singularity of the multi-domains leads to the discontinuity of the Hamiltonian. Under a controllability assumption on the fast variables, the limit equation (as the velocity of the fast variable goes to infinity) is obtained via a PDE approach and by means of the tools of the control theory.

5.3.5. *Optimal control of first order HJ equations with linearly bounded Hamiltonian*
Participant: Philip Graber.

In [40], we consider the optimal control of solutions of first order Hamilton-Jacobi equations, where the Hamiltonian is convex with linear growth. This models the problem of steering the propagation of a front by constructing an obstacle. We prove existence of minimizers to this optimization problem as in a relaxed setting and characterize the minimizers as weak solutions to a mean field game type system of coupled partial differential equations. Furthermore, we prove existence and partial uniqueness of weak solutions to the PDE system. An interpretation in terms of mean field games is also discussed.

5.3.6. *Zubov's equation for state-constrained perturbed nonlinear systems*
Participant: Hasnaa Zidani.

The paper [41] gives a characterization of the uniform robust domain of attraction for a nite non-linear controlled system subject to perturbations and state constraints. We extend the Zubov approach to characterize this domain by means of the value function of a suitable in nite horizon state-constrained control problem which at the same time is a Lyapunov function for the system. We provide associated Hamilton-Jacobi-Bellman equations and prove existence and uniqueness of the solutions of these generalized Zubov equations.

5.3.7. *Numerical methods for chance-constrained stochastic optimal control problems*
Participant: Laurent Pfeiffer.

In Laurent Pfeiffer's PhD, we study stochastic optimal control problems with a probability constraint on the final state. This constraint must be satisfied with a probability greater or equal than a given level. We analyse and compare two approaches for discrete-time problems: a first one based on a dynamic programming principle and a second one using Lagrange relaxation. These approaches can be used for continuous-time problems, for which we give numerical illustrations.

<p style="text-align:center">**DAHU Project-Team**</p>

5. New Results

5.1. Specification and Verification of Database Driven Systems

Participants: Serge Abiteboul, Luc Segoufin, Victor Vianu.

We continued our investication on the verification of database driven systems using an automata model with registers. We have exhibited new classes of decidable scenarios using nominal set theory [25]. These new classes contain the previously known relational cases but also the some semistructered ones.

We introduce in [24] and study a model of collaborative data-driven workflows. In a local-as-view style, each peer has a partial view of a global instance that remains purely virtual. Local updates have side effects on other peers' data, defined via the global instance. We also assume that the peers provide (an abstraction of) their specifications, so that each peer can actually see and reason on the specification of the entire system. We study the ability of a peer to carry out runtime reasoning about the global run of the system, and in particular about actions of other peers, based on its own local observations. A main contribution is to show that, under a reasonable restriction (namely, key-visibility), one can construct a finite symbolic representation of the infinite set of global runs consistent with given local observations. Using the symbolic representation, we show that we can evaluate in pspace a large class of properties over global runs, expressed in an extension of first-order logic with past linear-time temporal operators, PLTL-FO. We also provide a variant of the algorithm allowing to incrementally monitor a statically defined property, and then develop an extension allowing to monitor an infinite class of properties sharing the same temporal structure, defined dynamically as the run unfolds. Finally, we consider an extension of the language, augmeting work-flow control with PLTL-FO formulas. We prove that this does not increase the power of the workflow specification language, thereby showing that the language is closed under such introspective reasoning.

5.2. Distributed data management

Participants: Serge Abiteboul, Émilien Antoine, Cristina Sirangelo.

We have studied the feasibility of query answering in the presence of incomplete information in data. In particular we have investigated when it is the case that classical query evaluation techniques, which are commonly used over complete data, suffice to answer queries also in the presence of incompleteness [26]. These results allowed to find syntactic classes of queries that can be answered efficiently under many well known semantics of incompleteness, using query answering techniques which are already implemented (and optimized) in classical database systems.

The management of Web users' personal information is increasingly distributed across a broad array of applications and systems, including online social networks and cloud-based services. While users wish to share and integrate data using these systems, it is increasingly difficult to avoid the risks of unintended disclosures or unauthorized access by applications.

In [21], [20], we propose a novel access control model that operates within a distributed data management framework based on datalog. Using this model, users can control access to data they own and control applications they run. They can conveniently specify access control policies providing flexible tuple-level control derived using provenance information. We present a formal specification of the model, a theoretical analysis, and an implementation. We show that the computational cost of access control is acceptable.

5.3. Query Processing for the Web

Participants: Johann Brault-Baron, Arnaud Durand, Nadime Francis, Wojciech Kazana, Luc Segoufin, Cristina Sirangelo.

In many applications the output of a query may have a huge size and enumerating all the answers may already consume too many of the allowed resources. In this case it may be appropriate to first output a small subset of the answers and then, on demand, output a subsequent small numbers of answers and so on until all possible answers have been exhausted. To make this even more attractive it is preferable to be able to minimize the time necessary to output the first answers and, from a given set of answers, also minimize the time necessary to output the next set of answers - this second time interval is known as the *delay*. We have shown that this was doable with a linear preprocessing time and constant enumeration delay for first-order queries over structures of bounded expansion [27] and for monadic second-order queries over structures of bounded tree-width [15]. We also presented a survey about this work at the Intl. Conf. on Database Theory (ICDT) [19].

Web data is often structured in the XML format. In [18] we have surveyed results about static analysis of pattern-based queries over XML documents. These queries are analogs of conjunctive queries, their unions and Boolean combinations, in which tree patterns play the role of atomic formulae. These can be viewed as both queries and incomplete documents, and thus static analysis problems can also be viewed as answering queries over such documents. We looked at satisfiability of patterns under schemas, containment of queries for various features of XML used in queries, query answering, and applications of pattern-based queries in reasoning about schema mappings for data exchange.

<div align="center">

DEFI Project-Team

</div>

6. New Results

6.1. Qualitative methods for inverse scattering problems

6.1.1. *A generalized formulation of the Linear Sampling Method*
Participants: Lorenzo Audibert, Houssem Haddar.

We proposed and analyzed a new formulation of the Linear Sampling Method that uses an exact characterization of the targets shape in terms of the so-called farfield operator (at a fixed frequency). This characterization is based on constructing nearby solutions of the farfield equation using minimizing sequences of a least squares cost functional with an appropriate penalty term. We first provided a general framework for the theoretical foundation of the method in the case of noise-free and noisy measurements operator. We then explicited applications for the case of inhomogeneous inclusions and indicate possible straightforward generalizations. We finally validated the method through some numerical tests and compare the performances with classical LSM and the factorization methods.

6.1.2. *Inverse problems for periodic penetrable media*
Participant: Dinh Liem Nguyen.

Imaging periodic penetrable scattering objects is of interest for non-destructive testing of photonic devices. The problem is motivated by the decreasing size of periodic structures in photonic devices, together with an increasing demand in fast non-destructive testing. In this project, we considered the problem of imaging a periodic penetrable structure from measurements of scattered electromagnetic waves. As a continuation of earlier work jointly with A. Lechleiter we considered an electromagnetic problem for transverse magnetic waves (previous work treats transverse electric fields), and also the full Maxwell equations. In both cases, we treat the direct problem by a volumetric integral equation approach and construct a Factorization method.

6.1.3. *Transmission Eigenvalues and their application to the identification problem*
Participant: Houssem Haddar.

The so-called interior transmission problem plays an important role in the study of inverse scattering problems from (anisotropic) inhomogeneities. Solutions to this problem associated with singular sources can be used for instance to establish uniqueness for the imaging of anisotropic inclusions from muti-static data at a fixed frequency. It is also well known that the injectivity of the far field operator used in sampling methods is related to the uniqueness of solutions to this problem. The frequencies for which this uniqueness fails are called transmission eigenvalues. We are currently developing approaches where these frequencies can be used in identifying (qualitative informations on) the medium properties. Our research on this topic is mainly done in the framework of the associate team ISIP http://www.cmap.polytechnique.fr/~defi/ISIP/isip.html with the University of Delaware. A review article on the state of art concerning the transmission eigenvalue problem has been written in collaboration with F. Cakoni. We also edited a spacial issue of the journal Inverse Problems dedicated to the use of these transmission eigenvalues in inverse problems http://iopscience.iop.org/0266-5611/29/10/100201/. Our recent contributions are the following:

- Together with A. Cossonnière we analyzed the Fredholm properties of the interior transmission problem for the cases where the index contrast changes sign outside the boundary by using a surface integral equation approach.
- With F. Cakoni and N. Chaulet we investigated the asymptotic behaviour of the first transmission eigenvalue of a thin coating with respect to the coating thickness.

6.1.4. The factorization method for inverse scattering problems

6.1.4.1. The factorization method for cracks with impedance boundary conditions
Participant: Houssem Haddar.

With Y. Boukari we used the Factorization method to retrieve the shape of cracks with impedance boundary conditions from farfields associated with incident plane waves at a fixed fre- quency. This work is an extension of the study initiated by Kirsch and Ritter [Inverse Problems, 16, pp. 89-105, 2000] where the case of sound soft cracks is considered. We address here the scalar problem and provide theoretical validation of the method when the impedance boundary conditions hold on both sides of the crack. We then deduce an inversion algorithm and present some validating numerical results in the case of simply and multiply connected cracks [5].

6.1.4.2. The factorization method for EIT with uncertain background
Participants: Giovanni Migliorati, Houssem Haddar.

We extended the Factorization Method for Electrical Impedance Tomography to the case of background featuring uncertainty. This work is based on our earlier algorithm for known but inhomogeneous backgrounds. We developed three methodologies to apply the Factorization Method to the more difficult case of piece-wise constant but uncertain background. The first one is based on a recovery of the background through an optimization scheme and is well adapted to relatively low dimensional random variables describing the background. The second one is based on a weighted combination of the indicator functions provided by the Factorization Method for different realiza- tions of the random variables describing the uncertain background. We show through numerical experiments that this procedure is well suited to the case where many real-izations of the measurement operators are available. The third strategy is a variant of the previous one when measurements for the inclusion-free background are available. In that case, a single pair of measurements is sufficient to achieve comparable accuracy to the deterministic case [15].

6.1.4.3. The factorization method for GIBC
Participants: Mathieu Chamaillard, Houssem Haddar.

With N. Chaulet, we studied the identification of some obstacle and some Generalized Impedance Boundary Conditions (GIBC) on the boundary of such obstacle from far field measurements generated by the scattering of harmonic incident waves. The GIBCs are approximate models for thin coatings, corrugated surfaces, rough surfaces or imperfectly conducting media.

We justified the use of the Factorization method to solve the inverse obstacle problem in the presence of GIBCs. This method gives a uniqueness proof as well as a fast algorithm to reconstruct the obstacle from the knowledge of the far field produced by incident plane waves for all the directions of incidence at a given frequency. We also provided some numerical reconstructions of obstacles for several impedance operators.

6.2. Iterative Methods for Non-linear Inverse Problems

6.2.1. Inverse medium problem for axisymmetric eddy current models

Participants: Houssem Haddar, Zixian Jiang, Kamel Riahi.

We continued our developments of shape optimization methods for inclusion detection in an axisymmetric eddy current model. This problem is motivated by non-destructive testing methodologies for steam generators. We finalized a joint work with A. Lechleiter on numerical methods for the solution of the direct problem in weighted Sobolev spaces using approriate Dirichlet-to-Neumann mappings to bound the computational domain. We are also finalized jointly with M. El Guedri the work on inverse solver using a regularized steepest descent method for the problem of identifying a magnetite deposits using axial eddy current probe. We then addressed two new issues:

- We developed asymptotic models to identify thin highly conducting deposits. We derived three possible asymptotic models that can be exploited in the inverse problem. The numerical validation is under study.

- We extended the inverse scheeme to 3D configurations with axisymmetry at infinity: this includes exact charactherization of the shape derivative for a mixed formulation of eddy current problems and a parametric inversion scheme based on a pre-defined discrete grid for deposit location.

6.2.2. *The conformal mapping method and inverse scattering at low frequencies*
Participant: Houssem Haddar.

Together with R. Kress we employed a conformal mapping technique for the inverse problem to reconstruct a perfectly conducting inclusion in a homogeneous background medium from Cauchy data for electrostatic imaging, that is, for solving an inverse boundary value problem for the Laplace equation. In a recent work [13] we proposed an extension of this approach to inverse obstacle scattering for time-harmonic waves, that is, to the solution of an inverse boundary value problem for the Helmholtz equation. The main idea is to use the conformal mapping algorithm in an iterative procedure to obtain Cauchy data for a Laplace problem from the given Cauchy data for the Helmholtz problem. We presented the foundations of the method together with a convergence result and exhibit the feasibility of the method via numerical examples.

6.2.3. *A steepest descent method for inverse electromagnetic scattering problems*
Participant: Houssem Haddar.

Together with N. Chaulet, we proposed the application of a non linear optimization techniques to solve the inverse scattering problems for the $3D$ Maxwell's equations with generalized impedance boundary conditions. We characterized the shape derivative in the case where the GIBC is defined by a second order surface operator. We then applied a boundary variation method based on a regularized steepest descent to solve the 3-D inverse problem with partial farfield data. The obtained numerical results demonstrated the possibility of identifying the shape of coated objects as well as the parameters of the coating in the $3D$ Maxwell case.

6.3. Shape and topology optimization

6.3.1. *Geometric shape optimization*
Participant: Grégoire Allaire.

With Ch. Dapogny and P. Frey, we propose a new approach for geometry and topology optimization of structures which benefits from an accurate description of shapes at each stage of the iterative process (by means of a mesh amenable for mechanical analyses) while retaining the whole versatility of the level set method when it comes to accounting for their evolution. The key ingredients of this method are two operators for switching from a meshed representation of a domain to an implicit one, and conversely; this notably brings into play an algorithm for generating the signed distance function to an arbitrary discrete domain, and a mesh generation algorithm for implicitly-defined geometries.

6.3.2. *Worst-case design shape optimization*
Participant: Grégoire Allaire.

with Ch. Dapogny, we propose a deterministic method for optimizing a structure with respect to its worst possible behavior when a "small" uncertainty exists over some of its features. The main idea is to linearize the considered cost function with respect to the uncertain parameters, then to consider the supremum function of the obtained linear approximation, which can be rewritten as a more classical function of the design, owing to standard adjoint techniques from optimal control theory. The resulting linearized worst-case objective function turns out to be the sum of the initial cost function and of a norm of an adjoint state function, which is dual with respect to the considered norm over perturbations.

6.3.3. *Multi-phase structural optimization*
Participant: Grégoire Allaire.

With Ch. Dapogny, G. Delgado and G. Michailidis, we consider the optimal distribution of several elastic materials in a fixed working domain. In order to optimize both the geometry and topology of the mixture we rely on the level set method for the description of the interfaces between the different phases. We discuss various approaches, based on Hadamard method of boundary variations, for computing shape derivatives which are the key ingredients for a steepest descent algorithm. The shape gradient obtained for a sharp interface involves jump of discontinuous quantities at the interface which are difficult to numerically evaluate. Therefore we suggest an alternative smoothed interface approach which yields more convenient shape derivatives. We rely on the signed distance function and we enforce a fixed width of the transition layer around the interface (a crucial property in order to avoid increasing "grey" regions of fictitious materials). It turns out that the optimization of a diffuse interface has its own interest in material science, for example to optimize functionally graded materials.

6.3.4. *Level-Set Method*
Participant: Olivier Pantz.

We have begin to work, with Gabriel Delagado, on a new level-set optimization method, based on a gradient method. The key idea consists in computing directly the derivative of the discretized cost functions. The main advantage is that it is usually more simple to implement than the standard approach (consisting in using a discretized version of the gradient of the cost function). Moreover, the results obtained are as good or even better than the one obtained in previous works. Nevertheless, this method has its drawbacks, since the cost function is only derivable almost everywhere (the zero level-set has to be transverse to the triangulation of the mesh). It follows that convergence toward the minimum by the gradient method is not granted. To overcome this problem, we intend to use a mix-formulation for the state function. Unfortunately, such a formulation, in the case of linear elasticity is quite difficult to obtain. We thus intend to begin with the simplest scalar case, for which a lot more hybrid formulations are available.

6.3.5. *Optimization of a sodium fast reactor core*
Participants: Grégoire Allaire, Olivier Pantz.

In collaboration with D. Schmidt, G. Allaire and E. Dombre, we apply the geometrical shape optimization method for the design of a SFR (Sodium Fast reactor) core in order to minimize a thermal counter-reaction known as the sodium void effect. In this kind of reactor, by increasing the temperature, the core may become liable to a strong increase of reactivity ρ, a key-parameter governing the chain-reaction at quasi-static states. We first use the 1 group energy diffusion model and give the generalization to the 2 groups energy equation. We then give some numerical results in the case of the 1 group energy equation. Note that the application of our method leads to some designs whose interfaces can be parametrized by very smooth curves which can stand very far from realistic designs. We don't explain here the method that it would be possible to use for recovering an operational design but there exists several penalization methods that could be employed to this end. This work was partially sponsored by EDF. Our results will be published in the proceedings of the CEMRACS'11, during which part of the results have been obtained.

6.4. Asymptotic Analysis

6.4.1. *Effective boundary conditions for thin periodic coatings*
Participants: Mathieu Chamaillard, Houssem Haddar.

This topic is the object of a collaboration with Patrick Joly and is a continuation of our earlier work on interface conditions done in the framework of the PhD thesis of Berangère Delourme. Th goal here is to derive effective conditions that model scattering from thin periodic coatings where the thickness and the periodicity are of the same length but very small compared to the wavelength. The originality of our work, compared to abundant literature is to consider the case of arbitrary geometry (2-D or 3-D) and to consider higher order approximate models. We formally derived third order effective conditions after exhibiting the full asymptotic expansion of the solution in terms of the periodicity length.

6.4.2. Homogenization of electrokinetic models in porous media
Participant: Grégoire Allaire.

With R. Brizzi, J.-F. Dufrêche, A. Mikelic and A. Piatnitski, we are interested in the homogenization (or upscaling) of a system of partial differential equations describing the non-ideal transport of a N-component electrolyte in a dilute Newtonian solvent through a rigid porous medium. Our work can be divided in two different contributions. First, in the case of an ideal model (for which the homogenized system was already known) we consider the various limits which can be obtained in the effective parameters when the ratio between the characteristic pore length and the Debye length is either small or large. Second, we studied the homogenization process in the non-ideal case, namely when considering the so-called mean spherical approximation (MSA) model which takes into account finite size ions and screening effects.

6.4.3. A new shell modeling modeling
Participant: Olivier Pantz.

Using a formal asymptotic expansion, we have proved with K. Trabelsi, that non-isotropic thin-structure could behave (when the thickness is small) like a shell combining both membrane and bending effects. It is the first time to our knowledge that such a model is derived. An article on this is currently under review.

6.4.4. A new Liouville type Rigidity Theorem
Participant: Olivier Pantz.

We have recently developed a new Liouville type Rigidity Theorem. Considering a cylindrical shaped solid, we prove that if the local area of the cross sections is preserved together with the length of the fibers, then the deformation is a combination of a planar deformation and a rigid motion. The results currently obtained are limited to regular deformations and we are currently working with B. Merlet to extend them. Nevertheless, we mainly focus on the case where the conditions imposed to the local area of the cross sections and the length of the fibers are only "almost" fulfilled. This will enable us to derive rigorously new non linear shell models combining both membrane and flexural effects that we have obtained using a formal approach.

6.5. Diffusion MRI
Participants: Jing-Rebecca Li, Houssem Haddar, Dang Van Nguyen, Hang Tuan Nguyen.

Diffusion Magnetic Resonance Imaging (DMRI) is a promising tool to obtain useful information on microscopic structure and has been extensively applied to biological tissues. In particular, we would like to focus on two applications:

- inferring from DMRI measurements changes in the cellular volume fraction occurring upon various physiological or pathological conditions.

  This application is one of the first to show the promise of DMRI because it can detect acute cerebral ischemia (cell swelling) on the basis of lower than normal apparent diffusion coefficient a few minutes after stroke.

- estimating the average cell size in the case of tumor imaging

  This application is useful as a diagnostic tool as well as a tool for the evaluation of tumor treatments.

For both of the above applications we approach the problem via the following steps:

- Construct reduced models of the multiple-compartment Bloch-Torrey partial differential equation (PDE) using homogenization methods.
- Invert the resulting reduced models for the biological parameters of interest: the cellular volume fraction in the first case, and the average distance between neighboring cells in the second case.

../../../../projets/defi/IMG/cylindersandspheres.jpg

Figure 1. Computational domain for simulating diffusion in cerebral gray matter.

../../../../projets/defi/IMG/spheresqueezecells3d.jpg

Figure 2. Computational domain for simulating tumor cells.

We obtained the following results.

- We generated fairly complicated meshes that can be used to simulate diffusion in cerebral gray matter. In the Finite Elements code, this required using the mesh generation software Salome, developed at the CEA Saclay. We are working on the problem of increasing the cellular volume fraction to a physically realistic level, which is difficult for the mesh generator because of the very small distances between the neurons. An article describing the Finite Elements code has been accepted by the Journal of Computational Physics, to be published in 2014. An article on a version of the code using Finite Volume discretization has been accepted by Physics in Medicine and Biology, published in 2013.

- We developed a reduce model of the DMRI signal using homogenization methods. Two articles on this topic have been submitted.

DISCO Project-Team

6. New Results

6.1. Equidimensional block-triangular representation of linear functional systems

Participant: Alban Quadrat.

In [30], it is shown that every linear functional system (e.g., PD systems, differential time-delay systems, difference systems) is equivalent to a linear functional system defined by an upper block-triangular matrix of functional operators: each diagonal block is respectively formed by a generating set of the elements of the system satisfying a purely i-codimensional system. Hence, the system can be integrated in cascade by successively solving (inhomogeneous) i-codimensional linear functional systems to get a Monge parametrization of its solution space [120]. The results are based on an explicit construction of the grade/purity filtration of the module associated with the linear functional system. This new approach does not use complex Grothendieck spectral sequence arguments as is done in the literature of modern algebra [86], [87]. To our knowledge, the algorithm obtained in [30] is the most efficient algorithm existing in the literature of non-commutative algebra. It was implemented in the PURITYFILTRATION package developed in Maple (see Section 5.6) and in the homalg package of GAP 4 (see Section 5.7). Classes of overdetermined/underdetermined linear systems of partial differential equations which cannot be directly integrated by Maple can be solved using the PURITYFILTRATION package.

6.2. Serre's reduction of linear functional systems and related problems

Participants: Alban Quadrat, Thomas Cluzeau [ENSIL, Univ. Limoges].

Given a linear multidimensional system (e.g., ordinary/partial differential systems, differential time-delay systems, difference systems), Serre's reduction aims at finding an equivalent linear multidimensional system which contains fewer equations and fewer unknowns. Finding Serre's reduction of a linear multidimensional system can generally simplify the study of structural properties and of different numerical analysis issues, and it can sometimes help solving the linear multidimensional system in closed form. The connection between Serre's reduction and the decomposition problem [94], which aims at finding an equivalent linear functional system which is defined by a block diagonal matrix of functional operators, is algorithmically studied in [41], [42]. Moreover, a characterization of isomorphic finitely presented modules in terms of certain inflations of their presentation matrices is obtained in [42]. This result yields a connection between a certain matrix completion problem and Serre's reduction [42].

6.3. Algorithmic study of linear PD systems and Stafford's theorems

Participants: Alban Quadrat, Daniel Robertz [Univ. Aachen].

In [121],[82], algorithmic versions of Stafford's results [124] (e.g., computation of unimodular elements, decomposition of modules, Serre's splitting-off theorem, Stafford's reduction, Bass' cancellation theorem, minimal number of generators) were obtained and implemented in the STAFFORD package [82]. In particular, we show how a determined/overdetermined linear system of partial differential equations with either polynomial, rational, formal power series or locally convergent power series coefficients is equivalent to a linear system of partial differential equations with at most two unknowns. This result is a large generalization of the cyclic vector theorem which plays an important role in the theory of linear ordinary differential equations.

6.4. Foundations of the behavioural approach

Participant: Alban Quadrat.

Within the algebraic analysis approach to behaviours [91], [113], in [34], we propose to consider a system not only as a behaviour $\mathrm{ext}_D^0(M, \mathcal{F})$ [107], where M is the finitely presented left D-module defined by the matrix defining the system and $\mathcal{F}$ the signal space, but as the set of all the $\mathrm{ext}_D^i(M, \mathcal{F})$'s, where $0 \leq i \leq n$, where n is the global dimension of D. In this new framework, using Yoneda product, the left D-homomorphims of M [94] and the internal symmetries of the behaviour $\mathrm{ext}_D^0(M, \mathcal{F})$ [94] are generalized to the full system $\{\mathrm{ext}_D^i(M, \mathcal{F})\}_{i=0,\dots,n}$ In particular, a system-theoretic interpretation of the Yoneda product is given.

In [117], we study the construction of a double complex leading to a Grothendieck spectral sequence converging to the obstructions $\mathrm{tor}_D^i(N, \mathcal{F})$'s for the existence of a chain of successive parametrizations starting with the behaviour $\mathrm{ext}_D^0(M, \mathcal{F})$, where N is the Auslander transpose of M. These obstructions $\mathrm{tor}_D^i(N, \mathcal{F})$ can be studied by means of a long process starting with the $\mathcal{F}$-obstructions $\mathrm{ext}_D^j(\mathrm{ext}_D^k(N, D), \mathcal{F})$'s for the solvability of certain inhomogeneous linear systems defined by the algebraic obstructions $\mathrm{ext}_D^k(N, D)$'s measuring how far M is for being a projective left D-module. Hence, the algebraic properties of the left D-module M, defining the behaviour $\mathrm{ext}_D^0(M, \mathcal{F})$, and the functional properties of the signal space $\mathcal{F}$ can be simultaneously used to study the obstructions for the existence of a chain of successive parametrizations starting with the behaviour $\mathrm{ext}_D^0(M, \mathcal{F})$. These results can be used to find again the different situations studied in the literature (e.g., cases of an injective or a flat left D-module $\mathcal{F}$). Finally, setting $\mathcal{F} = D$, the above results can be used to find again the characterization of the grade/purity filtration of M by means of a Grothendieck spectral sequence. See Section 6.1 and [86], [87], [30].

Within the algebraic analysis approach to behaviours [91], [113], in [116], we explain how the concept of inverse image of a finitely presented left D-module M, defining the behaviour $\mathrm{ext}_D^0(M, \mathcal{F})$ [107], can be used to study the problem of characterizing the restriction of the behaviour $\mathrm{ext}_D^0(M, \mathcal{F})$ to a non characteristic submanifold of $\mathbb{R}^n$. In particular, we detail the explict construction of inverse images of left D-modules for standard maps.

6.5. Boundary value problems for linear ordinary integro-differential equations

Participants: Alban Quadrat, Georg Regensburger.

In [35], we study algorithmic aspects of linear ordinary integro-differential operators with polynomial coefficients. Even though this algebra is not noetherian and has zero divisors, Bavula recently proved in [85] that it is coherent, which allows one to develop an algebraic systems theory. For an algorithmic approach to linear systems theory of integro-differential equations with boundary conditions, computing the kernel of matrices is a fundamental task. As a first step, we have to find annihilators, which is, in turn, related to polynomial solutions. We present an algorithmic approach for computing polynomial solutions and the index for a class of linear operators including integro-differential operators. A generating set for right annihilators can be constructed in terms of such polynomial solutions. For initial value problems, an involution of the algebra of integro-differential operators also allows us to compute left annihilators, which can be interpreted as compatibility conditions of integro-differential equations with boundary conditions. These results are implemented in MAPLE based on the `IntDiffOp` and `IntDiffOperations` packages.

6.6. Noncommutative geometry approach to infinite-dimensional systems

Participant: Alban Quadrat.

In [112], [111], [110], it was shown how the fractional representation approach to analysis and synthesis problems developed by Vidyasagar, Desoer, Callier, Francis, Zames..., could be recast into a modern algebraic analysis approach based on module theory (e.g., fractional ideals, algebraic lattices) and the theory of Banach algebras. This new approach successfully solved open questions in the literature. Basing ourselves on this new approach, we explain in [114], [115] why the non-commutative geometry developed by Alain Connes is a natural framework for the study of stabilizing problems of infinite-dimensional systems. Using the 1-dimensional quantized calculus developed in non-commutative geometry and results obtained in [112], [111], [110], we show that every stabilizable system and their stabilizing controllers naturally admit geometric

structures such as connections, curvatures, Chern classes, ... These results developed in [114], [115] are the first steps toward the use of the natural geometry of the stabilizable systems and their stabilizing controllers in the study of the important H_∞ and H_2-problems.

6.7. Stability analysis of fractional neutral systems with commensurate delays

Participants: Le Ha Vy Nguyen, Catherine Bonnet.

Fractional neutral systems with commensurate delays have chains of poles aymptotic to vertical lines. The case where the imaginary axis is an asymptotic one is interesting. Indeed, if the system has some chains of poles asymptotic to the imaginary axis, then the fact that all poles lie in the open left half-plane does not guarantee the H_∞-stability of the system.

This kind of systems was studied in [97], [104]. In [97], systems with single chains of poles asymptotic to the imaginary axis was considered and necessary and sufficient conditions for H_∞-stability were derived. Some particular systems with multiple chains have been examined in [104]. We have extended this year this study to more general systems with multiple chains of poles approaching the imaginary axis.

6.8. Stabilization of fractional neutral systems with commensurate delays

Participants: Le Ha Vy Nguyen, Catherine Bonnet.

We consider fractional neutral systems with commensurate delays which may have chains of poles asymptotic to vertical lines lying in the open left half-plane and have chains clustering the imaginary axis. Due to the latter, the system may possess infinitely many poles in the right half-plane. We prove that a class of rational fractional controllers cannot stabilize this kind of systems in the sense of H_∞ except in a simple case. For this case, thanks to the fractional PI controller given in [1], a parametrization of stabilizing controllers is derived [105].

6.9. Stabilization of MISO fractional systems with delays

Participants: Le Ha Vy Nguyen, Catherine Bonnet.

In order to yield the set of all the stabilizing controllers of a class of MISO fractional systems with delays by mean of Youla-Kucera parametrization regarding H_∞-stability, we are interested in determining coprime factorizations of the transfer function. Explicit expressions of left coprime factorizations and left Bézout factors have been derived in [103]. We have continued this year to search for explicit expressions of right coprime factorizations for some classes of systems [63].

6.10. Interval Observer

Participants: Frédéric Mazenc [correspondent], Thach Ngoc Dinh, Silviu Iulian Niculescu.

We made several progresses in the domain of the construction of state estimators called interval observers.

1) In [18], we have shown how interval observers can be constructed for nonlinear (and not Lipschitz) systems possessing a special triangular system.

2) The contributions [20] and [55] present a new major result for the design of interval observers for discete-time systems with input and output: it is explained how two classical Luenberger obsevers can be used, even in the absence of the positivity property as interval observer, provided two appropriate output, which compose the lower and the upper bound of the interval observer, are selected. In [19], coordinate transformations which change an arbitrary linear discrete-time system into a positive one and general nonlinear design of interval observers for nonlinear systems (satisfying a restrictive stability assumption) are proposed.

3) The paper [54] presents the first construction of continuous-discrete interval observer for linear continuous-time systems with discrete measurements. The importance in engineering applications of this result is clear: most of the time the measured variables are available at discrete instants only. The result relies on the design of changes of coordinates which transform a linear system into a nonnegative one, but the dynamic part of interval observers is not cooperative.

6.11. Reduction model approach: new advances

Participants: Frédéric Mazenc [correspondent], Michael Malisoff [Louisiana State University], Silviu Iulian Niculescu, Dorothé Normand-Cyrot [L2S, CNRS].

We solved several distinct problems entailing to the celebrated reduction model approach. Let us recall that this technique makes it possible to stabilize systems with arbitrarily large pointwise or distributed delay in the input.

1) We proposed in [25] a new construction of exponentially stabilizing sampled feedbacks for continuous-time linear time-invariant systems with an arbitrarily large constant pointwise delay in the inputs. Stability is guaranteed under an assumption on the size of the largest sampling interval. The proposed design is based on an adaptation of the reduction model approach. The stability of the closed loop systems is proved through a Lyapunov-Krasovskii functional of a new type, from which is derived a robustness result

2) The paper [59] presents several results pertaining to the stabilization with feedbacks given by an explicit formula of linear time varying systems in the case where there is a constant delay in the input. In addition, it establishes input-to-state stability with respect to additive uncertainties. As a particular case, we considered a large class of rapidly time varying systems and provided a lower bound on the admissible rapidness parameters. We illustrated our results using a pendulum model.

3) The paper [24], which is devoted to the original problem of stabilizing nonlinear systems with input with distributed delay, is actually not an extension of the reduction model approach, but it complements it and uses operators which have been inspired by those used in the classical context of the reduction model theory.

6.12. Neutral systems and integral equations

Participants: Frédéric Mazenc [correspondent], Hiroshi Ito [Kyushu Institute of Technology], Pierdomenico Pepe [Univ. of L'Aquila].

1) For nonlinear systems with delay of neutral type, we developed a new technique of stability and robustness analysis. It relies on the construction of functionals which make it possible to establish estimates of the solutions different from, but very similar to, estimates of ISS or iISS type. These functionals are themselves different from, but very similar to, ISS or iISS Lyapunov-Krasovskii functionals. The approach applies to systems which do not have a globally Lipschitz vector field and are not necessarily locally exponentially stable. We apply this technique to carry out a backstepping design of stabilizing control laws for a family of neutral nonlinear systems [21].

2) In a second paper [57], we extended the previous results to the problem of deriving stability and stabilizability conditions for nonlinear systems with delay interconnected with an integral equation via the construction of a Lyapunov-Krasovskii functional.

6.13. Nonlinear systems with delay

Participants: Frédéric Mazenc [correspondent], Michael Malisoff [Louisiana State University], Thach Ngoc Dinh.

We obtained new results on the robustness analysis of nonlinear systems belonging to a general family when they are globally stabilized by a state feedback corrupted by the presence of a delay and sampling [22], [58]. The result is based on the construction of a non-quadratic Lyapunov-Krasovskii functional.

In [23], a problem of state feedback stabilization of time-varying feedforward systems with a pointwise delay in the input is solved. The approach we adopted relies on a time-varying change of coordinates and Lyapunov-Krasovskii functionals. The result applies for any given constant delay, and provides uniformly globally asymptotically stabilizing controllers of arbitrarily small amplitude. The closed-loop systems enjoy input-to-state stability properties with respect to additive uncertainty on the controllers. The work was illustrated through a tracking problem for a model for high level formation flight of unmanned air vehicles.

6.14. Set theoretic fault detection and isolation

Participant: Sorin Olaru.

Fault-tolerant control theory is a well-studied topic but the use of the sets in detection, isolation and/or reconfiguration is rather tangential. Sorin Olaru together with his collaborators (and principally with F. Stoican) conducted a systematic analysis of the set-theoretic elements and devise approaches which exploit advanced elements within the field. The main idea is to translate fault detection and isolation conditions into those conditions involving sets. Furthermore, these are to be computed efficiently using positive invariance and reachability notions. Constraints imposed by exact fault control are used to define feasible references (which impose persistent excitation and, thus, non-convex feasible sets). Particular attention is given to the reciprocal influences between fault detection and isolation on the one hand, and control reconfiguration on the other. The recent results on this topic are gathered in the recent book [81].

A new result has been obtained by the use of controlled invariance for the separation of faulty/healthy invariant sets in the detection and isolation [32] based on the necessary and sufficent conditions of George Bitsoris.

In a series of recent papers [67], [68], [70], [69], the link between the interval observers and the invariant sets have been investigated by establishing a series of formal results on the limit behaviour with potential applications in the detection and isolation of actuators faults

6.15. Model Predictive Control: distributed formulations and collision avoidance problems

Participant: Sorin Olaru.

In [78], the mixt integer techniques have been analysed in the distributed model predictive control context, underlining the dependence of collision avoidance mechanism on the obstacle modeling and susequently on their treatement inside optimization-based control techniques as MPC (model predictive control). On the same topic of adversary consraints, a geometrical conditions has been establised in [71] for the local stabilization of a linear dynamics on a boundary of a forbidden region in the state space.

The theoretical developments from the last two years on the MPC design for multi-agent control problem led to the succesful application of receding horizon flight control for trajectory tracking of autonomous aerial vehicles [28]. In the same line or research, the predictive control for trajectory tracking and decentralized navigation of multi-agent formations has been presented in [29].

In [66] a Characterization of the Relative Positioning of Mobile Agents for Full Sensorial Coverage in an Augmented Space with Obstacles is presented in view of a MPC control design.

A predictive control-based algorithm for path following of autonomous aerial vehicles has been proposed in [65] to improve the previous trajectory tracking mechanism. The ultimate goal oof both schemes is to avoid the real-time infeasibility problems in MPC.

The distributed predictive control mechanisms have been used for the control of a four interconnected tanks benchmark [48], proving the versatility of a nonlinear Distributed MPC technique previously proposed by A. Grancharova.

In [62] the distributed Model Predictive Control of Leader-Follower Systems has been studied using an interior point method with efficient computations leading to simple tuning mechanisms for the cost functions and the terminal sets of the local MPC sub-problems.

6.16. Invariant sets in control

Participant: Sorin Olaru.

The longstanding research interest on the positive invaraince of a set with respect to the trajectories of the dynamical systems allowed recently the statement of explicit invariant approximation of the maximal robust positive invariant ser for LTI dynamics with zonotopic disturbances [51].

In the class of hybrid dynamical systems, explicit robustness and fragility margins for discrete-time linear systems with PWA control has been established in [64] by means of positive invariance arguments.

In [37] a series of new results on the linear constrained regulation problem have been presented by completing the classical results with the case of active constraints for the equilibrium point.

6.17. Optimization of mu-analysis parameterization

Participant: Guillaume Sandou [correspondent].

The robustness against parametric uncertainties can be studied using the structured singular value mu. In that case, a normalization of the uncertain parameters is performed, and the mu analysis provides the larger parallelepiped centered in the nominal and included in the stability domain. However, results depend on the initial normalization. In this study, the normalization is optimized so as to get the largest guaranteed stability domain. The corresponding problem being highly nonlinear, a metaheuristic method, Particle Swarm Optimization, is used for that purpose. An academic and a real life example, namely the pendulum in the cart problem, have been used to prove the viability of the approach.

6.18. Optimal weight tuning in Hinfinity loop-shaping with PSO considering time constraints

Participants: Guillaume Sandou [correspondent], Gilles Duc [Supélec, E3S], Philippe Feyel [Sagem].

Hinfinity loop-shaping controllers have proven their efficiency to solve problems based on complex industrial specifications. However, the tuning of the weighting filters is a time consuming task. This work deals with the use of metaheuristics optimization for this weighting filter tuning. Whereas this topic has already been investigated in lot of works, all of them assume a particular pole/zero/damping/pulse expression for the searched transfer function. But choosing the best weight structure is not trivial and may lead to suboptimal solutions for the design process. That is why, we propose to enhance the weight selection problem by relaxing the structure constraints of transfer functions. The developed methodology is tested, using a real industrial example and leads to satisfactory results.

6.19. mu-synthesis with dynamic D-Scalings using Quantum Particle Swarm Optimization

Participants: Guillaume Sandou [correspondent], Gilles Duc [Supélec, E3S], Philippe Feyel [Sagem].

This study proposes to revisit the mu-synthesis problem with a recent and efficient meta-heuristic called Quantum Particle Swarm Optimization (QPSO). This algorithm allows us to optimizing dynamics (or static) D-scalings without fitting them which leads to robust performance controllers. This method has been applied to an industrial problem and has been proven to be better than the classical D-K iteration method.

6.20. Stabilization of time-delay systems

Participants: Alban Quadrat, Arnaud Quadrat [SAGEM, MASSY].

In [118], [122], we study the stabilization problem of a linear system formed by a simple integrator and a time-delay system. We show that the stabilizing controllers of such a system can be rewritten as the closed-loop system defined by the stabilizing controllers of the simple integrator and a distributed delay system. This result is used to study tracking problems appearing in the study of inertially stabilized platforms for optical imaging systems. Moreover, an elementary proof for the parametrization [111] of all stabilizing controllers of a stabilizable plant − which does not necessarily admits doubly coprime factorizations − is given in [122].

6.21. A Stabilization problem in chemostats

Participants: Frédéric Mazenc [correspondent], Jérôme Harmand [LBE INRA, EPI MODEMIC].

We have considered the classical model of the chemostat (which is a bio-reactor) with one substrate, one species and a Haldane type growth rate function is considered. The input substrate concentration is supposed to be constant and the dilution rate is considered as the control. The problem of globally asymptotically stabilizing a positive equilibrium point of this system in the case where the measured concentrations are delayed and piecewise constant with a piecewise constant control is addressed. The result relies on the introduction of a dynamic extension of a new type. [56].

6.22. Control design for UAVs

Participants: Frédéric Mazenc [correspondent], Michael Malisoff [Louisiana State University].

In [14], we studied a kinematic model that is suitable for control design for high level formation flight of UAVs (Unmanned Aerial Vehicles). We designed controllers that give robust global tracking for a wide class of reference trajectories in the sense of the robustness notion called input-to-state stability. The control laws satisfy amplitude and rate constraints.

6.23. Modeling and control of Acute Myeloid Leukemia

Participants: José Luis Avila Alonso [correspondent], Annabelle Ballesta [BANG project-team], Frédéric Bonnans [COMMANDS project-team], Catherine Bonnet, Jean Clairambault [BANG project-team], Xavier Dupuis [COMMANDS project-team], Pierre Hirsch [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Jean-Pierre Marie [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Faten Merhi [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Silviu Iulian Niculescu, Hitay Özbay [Bilkent University, Ankara, Turkey], Ruoping Tang [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris].

In [75] we propose a new mathematical model of the cell dynamics in Acute Myeloid Leukemia (AML) which takes into account the four different phases of the proliferating compartment. The dynamics of the cell populations are governed by transport partial differential equations structured in age and by using the method of characteristics, we obtain that the dynamical system of equation can be reduced to two coupled nonlinear equations with four internal sub-systems involving distributed delays. Local stability conditions for a particular equilibrium point, corresponding to a positive cells, are derived in terms of a set of inequalities involving the parameters of the mathematical model. A parameter estimation of our model is being performed using biological data (Annabelle Ballesta).

We have also studied a coupled model for healthy and cancer cell dynamics in Acute Myeloid Leukemia consisting of two stages of maturation for cancer cells and three stages of maturation for healthy cells. The cell dynamics are modelled by nonlinear partial differential equations. The interconnection phenomenon between the healthy and cancer cells takes place on the re-introduction functions leaving the resting compartments to the proliferating compartments of both populations of cells at the first stage. For a particular healthy equilibrium point, locally stability conditions involving the parameters of the mathematical model are obtained [83], [84].

GALEN Project-Team

6. New Results

6.1. Shape, Grouping and Recognition

6.1.1. Descriptors

Participants: Eduard Trulls, Iasonas Kokkinos.

In [30] we have extended our prior work on dense scale- and rotation- invariant image descriptors to take into account soft segmentation information. This allows us to discard measurements stemming from background structures, and as such renders our descriptors invariant to background changes and occlusions. This has allowed us to obtain state-of-the-art results on tasks such as large-displacement optical flow and wide-baseline stereo. We have made the implementation of these descriptors publicly available.

6.1.2. 3D structure detection

Participants: Haithem Boussaid, Iasonas Kokkinos.

In [22] we have started exploring the potential of combinatorial optimization in the medical imaging realm. We cast the problem of finding a 3D structure (a brain tumor) as that of finding the mode of a nonparametric distribution, constructed through Kernel Density Estimation. Current techniques for doing this (e.g. Mean Shift mode-seeking, Fast Gauss Transforms, etc.) are either iterative, or linear in the number of pixels, with a typically large constant. Instead, we develop a scheme that involves a very low-constant linear-time preprocessing step, and then uses Branch-and-Bound for fast mode estimation. As such it is scalable to large volumes, and serves as a rapid initialization of a region segmentation algorithm.

6.1.3. Facade parsing

Participants: Olivier Teboul, Iasonas Kokkinos, Loic Simon, Panagiotis Katsourakis, Nikos Paragios.

In [17] we pursue a Reinforcement Learning-based approach to couple image observations with a grammar-based method to partitioning a building facade. For this we expressed 2D grammar- based image parsing as a Markov decision process where an agent has to take actions in an environment so as to maximize some notion of cumulative reward (reflecting the segmentation quality). This allowed us to accelerate previous stochastic hill-climbing approaches to image parsing by more than an order of magnitude.

6.1.4. Fast object detection

Participant: Iasonas Kokkinos.

In [27] we extended our previous work on fast object detection by developing a sparse-coding method for the efficient sharing of computation among multiple object models. In particular the first processing step of 'part score' computation was originally performed separate per object category; instead, we propose to do it 'in batch mode', so as to exploit the commonalities that exist among object parts. Building on recent developments in sparse coding we have managed to construct a compact basis for this task, which in the end gave us a two-fold acceleration over our previous fastest algorithms.

6.2. Machine Learning

6.2.1. Discriminative Parameter Estimation for Random Walks Segmentation

Participants: Pierre-Yves Baudin, Puneet Kumar, Noura Azzabou, Pierre Carlier, Nikos Paragios, M. Pawan Kumar Blaschko

In [19], we proposed a a novel discriminative learning framework that estimates the parameters of a random walks segmentation framework using a training dataset. The main challenge we face is that the training samples are not fully supervised. Specifically, they provide a hard segmentation of the medical images, instead of a probabilistic segmentation. We overcome this challenge by treating the optimal probabilistic segmentation that is compatible with the given hard segmentation as a latent variable. This allows us to employ the latent support vector machine (LSVM) formulation for parameter estimation.

6.2.2. *Structured Sparsity & Applications*

Participants: Katerina Gkirtzou, Wojciech Zaremba, Matthew Blaschko, M. Pawan Kumar, Nikos Paragios

We developed several machine learning applications to fMRI data, including graph representations [25] and structured sparsity regularization [26], [44]. A similar structured sparsity approach was applied in the development of a novel learning algorithm, the k-support regularized SVM, with applications to neuromuscular disease classification from diffusion tensor imaging [24]. Efficient training applications for taxonomic classification were developed in [21], while a fine grained taxonomic image classification task was introduced in [45]. The role of non-maximal suppression in accurate and efficient object detection cascades was elucidated in [20]. A fast, consistent two-sample test based on kernelized statistics was developed in [33].

6.2.3. *Learning from M/EEG Data with Variable Brain Activation Delays*

Participants: Wojciech Zaremba, Alexander Gramfort, M. Pawan Kumar, Matthew Blaschko

In [34], propose to address the misalignment of M/EEG samples by explicitly modeling time shifts of different brain responses in a classification setup. To this end, we use the LSVM formulation, where the latent shifts are inferred while learning the classifier parameters. The inferred shifts are further used to improve the signal-to-noise ratio of the M/EEG data, and to infer the chronometry and the sequence of activations across the brain regions that are involved in the experimental task.

6.3. Biomedical Image Analysis

6.3.1. *Reconstruction*

Participants: Helen Langet, Nikos Paragios

In [38] an overview of the methodological foundations of biomedical image analysis as well as their use to provide answers to a variety of clinical problems are presented. The problem of volumes of rotational angiography using non-linear sparsity constraints was studied in [28] where a novel method able to handle highly under-sampled acquisitions was introduced.

6.3.2. *Graphical models and Image Segmentation*

Participants: Bo Xiang, Nikos Paragios

[18] presents an overview of the use of graphical models in artificial vision where both inference, learning as well as applications are discussed. In [32] a max-margin dual decomposition method was used towards learning the compact, pose invariant shape representation using higher order graphs acting both on the connectivity of the graph as well its potentials. Graphical model was used as prior in [13] under a "curve" propagation principle for generic prior-constrained organ segmentation in 2D images. Similar inspiration driven from a higher order pose invariant graphical model learned according to [32] was considered in [31] where a novel segmentation method was proposed coupling model-based and pixel-based concepts while being pose invariant. The underlying idea was to consider a two-layer interconnected graphical model acting on pixel and on control points where segmentation consistency was imposed throgh penalties on label discrepancies of the different layers. Higher order graphical models were also employed in [14] for spine segmentation using an articulated graphical model where a non-linear approach/embedding towards reducing the complexity of the inference step was considered at training.

6.3.3. *Deformable Registration and Fusion*

Participants: Enzo Ferrante, Sarah Parisot, Nikos Paragios

In [16] a comprehensive survey of deformable registration was presented. It was organized in three sections: the first was studying the deformation model, the second the similarity criterion while the last section discussed the different optimization strategies. The problem of atlas-based segmentation/registration in the presence of brain tumors was studied in [29] an adaptive uncertainty-driven sampling strategy was proposed coupling segmentation and registration. Both sampling spaces (quantization of the search space, deformation grid) were determined according to the observed optimization min-marginals. The challenging problem of image to slice registration was proposed in [23] where an over-parameterized low rank graphical model acting both on the plan selection as well the in-plane deformations was introduced. The main strength of the method was its ability to simultaneously recover both the plane and the organ deformation.

GECO Project-Team

6. New Results

6.1. New results: geometric control

We start by presenting some results on motion planning and tracking algorithms.

- In [22] we study the complexity of the motion planning problem for control-affine systems. Such complexities are already defined and rather well-understood in the particular case of nonholonomic (or sub-Riemannian) systems. Our aim is to generalize these notions and results to systems with a drift. Accordingly, we present various definitions of complexity, as functions of the curve that is approximated, and of the precision of the approximation. Due to the lack of time-rescaling invariance of these systems, we consider geometric and parametrized curves separately. Then, we give some asymptotic estimates for these quantities.

- In [23] we study the problem of controlling an unmanned aerial vehicle (UAV) to provide a target supervision and to provide convoy protection to ground vehicles. We first present a control strategy based upon a Lyapunov–LaSalle stabilization method to provide supervision of a stationary target. The UAV is expected to join a pre-designed admissible circular trajectory around the target which is itself a fixed point in the space. Our strategy is presented for both HALE (High Altitude Long Endurance) and MALE (Medium Altitude Long Endurance) types UAVs. A UAV flying at a constant altitude (HALE type) is modeled as a Dubins vehicle (i.e. a planar vehicle with constrained turning radius and constant forward velocity). For a UAV that might change its altitude (MALE type), we use the general kinematic model of a rigid body evolving in $\mathbb{R}^3$. Both control strategies presented are smooth and unlike what is usually proposed in the literature these strategies asymptotically track a circular trajectory of exact minimum turning radius. We then consider the problem of adding to the tracking task an optimality criterion. In particular, we present the time-optimal control synthesis for tracking a circle by a Dubins vehicle. This optimal strategy, although much simpler than the point-to-point time-optimal strategy obtained by P. Souéres and J.-P. Laumond in the 1990s, is very rich. Finally, we propose control strategies to provide supervision of a moving target, that are based upon the previous ones.

- In [26] we prove the continuity and the Hölder equivalence w.r.t. an Euclidean distance of the value function associated with the L^1 cost of the control-affine system $\dot{q} = f_0(q) + \sum_{j=1}^m u_j f_j(q)$, satisfying the strong Hörmander condition. This is done by proving a result in the same spirit as the Ball-Box theorem for driftless (or sub-Riemannian) systems. The techniques used are based on a reduction of the control-affine system to a linear but time-dependent one, for which we are able to define a generalization of the nilpotent approximation and through which we derive estimates for the shape of the reachable sets. Finally, we also prove the continuity of the value function associated with the L^1 cost of time-dependent systems of the form $\dot{q} = \sum_{j=1}^m u_j f_j^t(q)$.

Let us list some new results in sub-Riemannian geometry and hypoelliptic diffusion.

- In [1] we provide normal forms for 2D almost-Riemannian structures, which are generalized Riemannian structures on surfaces for which a local orthonormal frame is given by a Lie bracket generating pair of vector fields that can become collinear. Generically, there are three types of points: Riemannian points where the two vector fields are linearly independent, Grushin points where the two vector fields are collinear but their Lie bracket is not, and tangency points where the two vector fields and their Lie bracket are collinear and the missing direction is obtained with one more bracket. We consider the problem of finding normal forms and functional invariants at each type of point. We also require that functional invariants are "complete" in the sense that they permit to recognize locally isometric structures. The problem happens to be equivalent to the one of finding a smooth canonical parameterized curve passing through the point and being transversal to the distribution.

For Riemannian points such that the gradient of the Gaussian curvature K is different from zero, we use the level set of K as support of the parameterized curve. For Riemannian points such that the gradient of the curvature vanishes (and under additional generic conditions), we use a curve which is found by looking for crests and valleys of the curvature. For Grushin points we use the set where the vector fields are parallel. Tangency points are the most complicated to deal with. The cut locus from the tangency point is not a good candidate as canonical parameterized curve since it is known to be non-smooth. Thus, we analyze the cut locus from the singular set and we prove that it is not smooth either. A good candidate appears to be a curve which is found by looking for crests and valleys of the Gaussian curvature. We prove that the support of such a curve is uniquely determined and has a canonical parametrization.

- The curvature discussed in [14] is a rather far going generalization of the Riemann sectional curvature. We define it for a wide class of optimal control problems: a unified framework including geometric structures such as Riemannian, sub-Riemannian, Finsler and sub-Finsler structures; a special attention is paid to the sub-Riemannian (or Carnot–Caratheodory) metric spaces. Our construction of the curvature is direct and naive, and it is similar to the original approach of Riemann. Surprisingly, it works in a very general setting and, in particular, for all sub-Riemannian spaces.

- In [15] we provide the small-time heat kernel asymptotics at the cut locus in three relevant cases: generic low-dimensional Riemannian manifolds, generic 3D contact sub-Riemannian manifolds (close to the starting point) and generic 4D quasi-contact sub-Riemannian manifolds (close to a generic starting point). As a byproduct, we show that, for generic low-dimensional Riemannian manifolds, the only singularities of the exponential map, as a Lagragian map, that can arise along a minimizing geodesic are A_3 and A_5 (in the classification of Arnol'd's school). We show that in the non-generic case, a cornucopia of asymptotics can occur, even for Riemannian surfaces.

- In [19] we study the evolution of the heat and of a free quantum particle (described by the Schroedinger equation) on two-dimensional manifolds endowed with the degenerate Riemannian metric $ds^2 = dx^2 + |x|^{-2\alpha} d\theta^2$, where $x \in \mathbb{R}$, $\theta \in \mathbb{T}$ and the parameter $\alpha \in \mathbb{R}$. For $\alpha \leq -1$ this metric describes cone-like manifolds (for $\alpha = -1$ it is a flat cone). For $\alpha = 0$ it is a cylinder. For $\alpha \geq 1$ it is a Grushin-like metric. We show that the Laplace–Beltrami operator Δ is essentially self-adjoint if and only if $\alpha \notin (-3, 1)$. In this case the only self-adjoint extension is the Friedrichs extension Δ_F, that does not allow communication through the singular set $\{x = 0\}$ both for the heat and for a quantum particle. For $\alpha \in (-3, -1]$ we show that for the Schroedinger equation only the average on θ of the wave function can cross the singular set, while the solutions of the only Markovian extension of the heat equation (which indeed is Δ_F) cannot. For $\alpha \in (-1, 1)$ we prove that there exists a canonical self-adjoint extension Δ_B, called bridging extension, which is Markovian and allows the complete communication through the singularity (both of the heat and of a quantum particle). Also, we study the stochastic completeness (i.e., conservation of the L^1 norm for the heat equation) of the Markovian extensions Δ_F and Δ_B, proving that Δ_F is stochastically complete at the singularity if and only if $\alpha \leq -1$, while Δ_B is always stochastically complete at the singularity.

6.2. New results: quantum control

New results have been obtained for the control of the bilinear Schrödinger equation.

- In [4] we show the approximate rotational controllability of a polar linear molecule by means of three nonresonant linear polarized laser fields. The result is based on a general approximate controllability result for the bilinear Schroedinger equation, with wavefunction varying in the unit sphere of an infinite-dimensional Hilbert space and with several control potentials, under the assumption that the internal Hamiltonian has discrete spectrum. A further general results, extending the above approach, are obtained in [16].

- In [5] we provide a short introduction to modern issues in the control of infinite dimensional closed quantum systems, driven by the bilinear Schroedinger equation. The first part is a quick presentation

of some of the numerous recent developments in the fields. This short summary is intended to demonstrate the variety of tools and approaches used by various teams in the last decade. In a second part, we present four examples of bilinear closed quantum systems. These examples were extensively studied and may be used as a convenient and efficient test bench for new conjectures. Finally, we list some open questions, both of theoretical and practical interest.

- In [6] we study the so-called spin-boson system, namely a spin-1/2 particle in interaction with a distinguished mode of a quantized bosonic field. We control the system via an external field acting on the bosonic part. Applying geometric control techniques to the Galerkin approximation and using perturbation theory to guarantee non-resonance of the spectrum of the drift operator, we prove approximate controllability of the system, for almost every value of the interaction parameter.

- In [9] and [25] we investigate the controllability of a quantum electron trapped in a two-dimensional device. The problem is modeled by the Schroedinger equation in a bounded domain coupled to the Poisson equation for the electrical potential. The controller acts on the system through the boundary condition on the potential, on a part of the boundary modeling the gate. We prove that, generically with respect to the shape and boundary conditions on the gate, the device is controllable. In [25] We also consider control properties of a more realistic nonlinear version of the device, taking into account the self-consistent electrostatic Poisson potential.

- In [18] we study the controllability of a closed control-affine quantum system driven by two or more external fields. We provide a sufficient condition for controllability in terms of existence of conical intersections between eigenvalues of the Hamiltonian in dependence of the controls seen as parameters. Such spectral condition is structurally stable in the case of three controls or in the case of two controls when the Hamiltonian is real. The spectral condition appears naturally in the adiabatic control framework and yields approximate controllability in the infinite-dimensional case. In the finite-dimensional case it implies that the system is Lie-bracket generating when lifted to the group of unitary transformations, and in particular that it is exactly controllable. Hence, Lie algebraic conditions are deduced from purely spectral properties. Another contribution of [18] is the proof that approximate and exact controllability are equivalent properties for general finite-dimensional quantum systems.

6.3. New results: neurophysiology

- In recent papers models of the human locomotion by means of an optimal control problem have been proposed. In this paradigm, the trajectories are assumed to be solutions of an optimal control problem whose cost has to be determined. The purpose of [3] is to analyze the class of optimal control problems defined in this way. We prove strong convergence result for their solutions on the one hand for perturbations of the initial and final points (stability), and on the other hand for perturbations of the cost (robustness).

- [8] analyses a class of optimal control problems on geometric paths of the euclidean space, that is, curves parametrized by arc length. In the first part we deal with existence and robustness issues for such problems and we define the associated inverse optimal control problem. In the second part we discuss the inverse optimal control problem in the special case of planar trajectories and under additional assumptions. More precisely we define a criterion to restrict the study to a convenient class of costs based on the analysis of experimentally recorded trajectories. This method applies in particular to the case of human locomotion trajectories.

- The article [17] presents an algorithm implementing the theory of neurogeometry of vision, described by Jean Petitot in his book. We propose a new ingredient, namely working on the group of translations and discrete rotations $SE(2, N)$. We focus on the theoretical and numerical aspects of integration of an hypoelliptic diffusion equation on this group. Our main tool is the generalized Fourier transform. We provide a complete numerical algorithm, fully parallellizable.

6.4. New results: switched systems

- In [2] we study the control system $\dot{x} = Ax + \alpha(t)bu$ where the pair (A, b) is controllable, $x \in \mathbb{R}^2$, $u \in \mathbb{R}$ is a scalar control and the unknown signal $\alpha : \mathbb{R}_+ \to [0, 1]$ is (T, μ)-persistently exciting (PE), i.e., there exists $T \geq \mu > 0$ such that, for all $t \in \mathbb{R}_+$, $\int_t^{t+T} \alpha(s)ds \geq \mu$. We are interested in the stabilization problem of this system by a linear state feedback $u = -Kx$. In [2], we positively answer a question asked in [52] and prove the following: Assume that the class of (T, μ)-PE signals is restricted to those which are M-Lipschitzian, where $M > 0$ is a positive constant. Then, given any $C > 0$, there exists a linear state feedback $u = -Kx$ where K only depends on (A, b) and T, μ, M so that, for every M-Lipschitzian (T, μ)-PE signal, the rate of exponential decay of the time-varying system $\dot{x} = (A - \alpha(t)bK)x$ is greater than C.

  In [20] we consider a family of linear control systems $\dot{x} = Ax + \alpha Bu$ where α belongs to a given class of persistently exciting signals. We seek maximal α-uniform stabilisation and destabilisation by means of linear feedbacks $u = Kx$. We extend previous results obtained for bidimensional single-input linear control systems to the general case as follows: if the pair (A, B) verifies a certain Lie bracket generating condition, then the maximal rate of convergence of (A, B) is equal to the maximal rate of divergence of $(-A, -B)$. We also provide more precise results in the general single-input case, where the above result is obtained under the sole assumption of controllability of the pair (A, B).

  The paper [24] considers the stabilization to the origin of a persistently excited linear system by means of a linear state feedback, where we suppose that the feedback law is not applied instantaneously, but after a certain positive delay (not necessarily constant). The main result is that, under certain spectral hypotheses on the linear system, stabilization by means of a linear delayed feedback is indeed possible, generalizing a previous result already known for non-delayed feedback laws.

  Several problems and results related with persistent excitation and stabilization are discussed in the survey [11]. These problems and results deal with both finite- and infinite-dimensional systems.

- In [7] we consider several time-discretization algorithms for singularly perturbed switched systems. The algorithms correspond to different sampling times and the discretization procedure respects the splitting of each mode in fast and slow dynamics. We study whether such algorithms preserve the asymptotic or quadratic stability of the original continuous-time singularly perturbed switched system.

- In [10] we consider affine switched systems as perturbations of linear ones, the equilibria playing the role of perturbation parameters. We study the stability properties of an affine switched system under arbitrary switching, assuming that the corresponding linear system is uniformly exponentially stable. It turns out that the affine system admits a minimal invariant set Ω, whose properties we investigate. In the two-dimensional bi-switched case when both subsystems have non-real eigenvalues we are able to characterize Ω completely and to prove that all trajectories of the system converge to Ω. We also explore the behavior of minimal-time trajectories in Ω by constructing optimal syntheses.

- In [21] we give a collection of converse Lyapunov–Krasovskii theorems for uncertain retarded differential equations. We show that the existence of a weakly degenerate Lyapunov–Krasovskii functional is a necessary and sufficient condition for the global exponential stability of the linear retarded functional differential equations. This is carried out using the switched system transformation approach.

GEOMETRICA Project-Team

6. New Results

6.1. Mesh Generation and Geometry Processing

6.1.1. Splat-based Surface Reconstruction from Defect-Laden Point Sets.

Participant: Mariette Yvinec.

In collaboration with Pierre Alliez (EPI Titane), Ricard Campos (University of Girona), Raphael Garcia (University of Girona)

We introduce a method for surface reconstruction from point sets that is able to cope with noise and outliers. First, a splat-based representation is computed from the point set. A robust local 3D RANSAC-based procedure is used to filter the point set for outliers, then a local jet surface – a low-degree surface approximation – is fitted to the inliers. Second, we extract the reconstructed surface in the form of a surface triangle mesh through Delaunay refinement. The Delaunay refinement meshing approach requires computing intersections between line segment queries and the surface to be meshed. In the present case, intersection queries are solved from the set of splats through a 1D RANSAC procedure. [14].

6.1.2. Constructing Intrinsic Delaunay Triangulations of Submanifolds

Participants: Jean-Daniel Boissonnat, Ramsay Dyer.

In collaboration with Arijit Ghosh (Indian Statistical Institute)

We describe an algorithm to construct an intrinsic Delaunay triangulation of a smooth closed submanifold of Euclidean space [42]. Using results established in a companion paper on the stability of Delaunay triangulations on δ-generic point sets, we establish sampling criteria which ensure that the intrinsic Delaunay complex coincides with the restricted Delaunay complex and also with the recently introduced tangential Delaunay complex. The algorithm generates a point set that meets the required criteria while the tangential complex is being constructed. In this way the computation of geodesic distances is avoided, the runtime is only linearly dependent on the ambient dimension, and the Delaunay complexes are guaranteed to be triangulations of the manifold.

6.1.3. Delaunay Triangulation of Manifolds

Participants: Jean-Daniel Boissonnat, Ramsay Dyer.

In collaboration with Arijit Ghosh (Indian Statistical Institute)

We present an algorithmic framework for producing Delaunay triangulations of manifolds [44]. The input to the algorithm is a set of sample points together with coordinate patches indexed by those points. The transition functions between nearby coordinate patches are required to be bi-Lipschitz with a constant close to 1. The primary novelty of the framework is that it can accommodate abstract manifolds that are not presented as submanifolds of Euclidean space. The output is a manifold simplicial complex that is the Delaunay complex of a perturbed set of points on the manifold. The guarantee of a manifold output complex demands no smoothness requirement on the transition functions, beyond the bi-Lipschitz constraint. In the smooth setting, when the transition functions are defined by common coordinate charts, such as the exponential map on a Riemannian manifold, the output manifold is homeomorphic to the original manifold, when the sampling is sufficiently dense.

6.1.4. Anisotropic Delaunay Meshes of Surfaces

Participants: Jean-Daniel Boissonnat, Mariette Yvinec.

In collaboration with Jane Tournois (GeometryFactory) and Kan-Le Shi (Tsing Hua University)

Anisotropic simplicial meshes are triangulations with elements elongated along prescribed directions. Anisotropic meshes have been shown to be well suited for interpolation of functions or solving PDEs. They can also significantly enhance the accuracy of a surface representation. Given a surface S endowed with a metric tensor field, we propose a new approach to generate an anisotropic mesh that approximates S with elements shaped according to the metric field [13], [47]. The algorithm relies on the well-established concepts of restricted Delaunay triangulation and Delaunay refinement and comes with theoretical guarantees. The star of each vertex in the output mesh is Delaunay for the metric attached to this vertex. Each facet has a good aspect ratio with respect to the metric specified at any of its vertices. The algorithm is easy to implement. It can mesh various types of surfaces like implicit surfaces, polyhedra or isosurfaces in 3D images. It can handle complicated geometries and topologies, and very anisotropic metric fields.

6.2. Topological and Geometric Inference

6.2.1. *An Efficient Data Structure for Computing Persistent Cohomology*

Participants: Jean-Daniel Boissonnat, Clément Maria.

In collaboration with Tamal Dey (Ohio State University)

Persistent homology with coefficients in a field F coincides with the same for cohomology because of duality. We propose an implementation of a recently introduced algorithm for persistent cohomology that attaches annotation vectors with the simplices. We separate the representation of the simplicial complex from the representation of the cohomology groups, and introduce a new data structure for maintaining the annotation matrix, which is more compact and reduces substancially the amount of matrix operations. In addition, we propose a heuristic to further simplify the representation of the cohomology groups and improve both time and space complexities. The paper provides a theoretical analysis, as well as a detailed experimental study of our implementation and comparison with state-of-the-art software for persistent homology and cohomology [41], [29].

6.2.2. *Multi-Field Persistent Homology*

Participants: Jean-Daniel Boissonnat, Clément Maria.

In [46], we introduce the *multi-field persistence diagram* for the persistence homology of a filtered complex. It encodes compactly the *superimposition* of the persistence diagrams of the complex with several field coefficients, and provides a substancially more precise description of the topology of the filtered complex. Specifically, the multi-field persistence diagram encodes the Betti numbers of integral homology and the prime divisors of the torsion coefficients of the underlying shape. Moreover, it enjoys similar stability properties as the ones of standard persistence diagrams, with the appropriate notion of distance. These properties make the multi-field persistence diagram a useful tool in computational topology. The multi-field algorithms are, in practice, as fast as algorithms that compute persistent homology in a single field.

6.2.3. *Zigzag Zoology: Rips Zigzags for Homology Inference*

Participants: Steve Oudot, Donald Sheehy.

For points sampled near a compact set X, the persistence barcode of the Rips filtration built from the sample contains information about the homology of X as long as X satisfies some geometric assumptions. The Rips filtration is prohibitively large, however zigzag persistence can be used to keep the size linear. We present several species of Rips-like zigzags and compare them with respect to the signal-to-noise ratio, a measure of how well the underlying homology is represented in the persistence barcode relative to the noise in the barcode at the relevant scales. Some of these Rips-like zigzags have been available as part of the Dionysus library for several years while others are new. Interestingly, we show that some species of Rips zigzags will exhibit less noise than the (non-zigzag) Rips filtration itself. Thus, Rips zigzags can offer improvements in both size complexity and signal-to-noise ratio. Along the way, we develop new techniques for manipulating and comparing persistence barcodes from zigzag modules. In particular, we give methods for reversing arrows and removing spaces from a zigzag while controlling the changes occurring in its barcode. We also discuss

factoring zigzags and a kind of interleaving of two zigzags that allows their barcodes to be compared. These techniques were developed to provide our theoretical analysis of the signal-to-noise ratio of Rips-like zigzags, but they are of independent interest as they apply to zigzag modules generally [33].

6.2.4. *Efficient and Robust Topological Data Analysis on Metric Spaces*

Participants: Mickaël Buchet, Frédéric Chazal, Steve Oudot, Donald Sheehy.

We extend the notion of the distance to a measure from Euclidean space to probability measures on general metric spaces as a way to perform topological data analysis in a way that is robust to noise and outliers. We then give an efficient way to approximate the sub-level sets of this function by a union of metric balls and extend previous results on sparse Rips filtrations to this setting. This robust and efficient approach to topological data analysis is illustrated with several examples from an implementation [54].

6.2.5. *Noise-Adaptive Shape Reconstruction from Raw Point Sets*

Participant: David Cohen-Steiner.

In collaboration with Pierre Alliez (EPI Titane), Simon Giraudot (EPI Titane)

We propose a noise-adaptive shape reconstruction method specialized to smooth, closed hypersurfaces. Our algorithm takes as input a defect-laden point set with variable noise and outliers, and comprises three main steps. First, we compute a novel type of robust distance function to the data. As a robust distance function, its sublevel-sets have the correct homotopy type when the data is a sufficiently good sample of a regular shape. The new feature is a built-in scale selection mechanism that adapts to the local noise level, under the assumption that the inferred shape is a smooth submanifold of known dimension. Second, we estimate the sign and confidence of the function at a set of seed points, based on estimated crossing parities along the edges of a uniform random graph. That component is inspired by the classical MAXCUT relaxation, except that we only require a linear solve as opposed to an eigenvector computation. Third, we compute a signed implicit function through a random walker approach with soft constraints chosen as the most confident seed points computed in previous step. The resulting pipeline is scalable and offers excellent behavior for data exhibiting variable noise levels [19].

6.2.6. *Optimal Rates of Convergence for Persistence Diagrams in Topological Data Analysis*

Participants: Frédéric Chazal, Marc Glisse, Bertrand Michel.

In collaboration with Catherine Labruère (Université de Bourgogne).

Computational topology has recently known an important development toward data analysis, giving birth to the field of topological data analysis. Topological persistence, or persistent homology, appears as a fundamental tool in this field. In this paper [57] (to appear in proc. ICML 2014), we study topological persistence in general metric spaces, with a statistical approach. We show that the use of persistent homology can be naturally considered in general statistical frameworks and persistence diagrams can be used as statistics with interesting convergence properties. Some numerical experiments are performed in various contexts to illustrate our results.

6.2.7. *Bootstrap and Stochastic Convergence for Persistence Diagrams and Landscapes*

Participant: Frédéric Chazal.

In collaboration with B. Fasy (Tulane University), F. Lecci, A. Rinaldo, A. Singh, L. Wasserman (Carnegie Mellon University).

Persistent homology probes topological properties from point clouds and functions. By looking at multiple scales simultaneously, one can record the births and deaths of topological features as the scale varies. We can summarize the persistent homology with the persistence landscape, introduced by Bubenik, which converts a diagram into a well-behaved real-valued function. We investigate the statistical properties of landscapes, such as weak convergence of the average landscapes and convergence of the bootstrap. In addition, we introduce an alternate functional summary of persistent homology, which we call the silhouette, and derive an analogous statistical theory [55].

6.2.8. Gromov-Hausdorff Approximation of Metric Spaces with Linear Structure
Participant: Frédéric Chazal.

In collaboration with S. Jian (Tsinghua University).

In many real-world applications data come as discrete metric spaces sampled around 1-dimensional filamentary structures that can be seen as metric graphs. In this paper [58] we address the metric reconstruction problem of such filamentary structures from data sampled around them. We prove that they can be approximated, with respect to the Gromov-Hausdorff distance by well-chosen Reeb graphs (and some of their variants) and we provide an efficient and easy to implement algorithm to compute such approximations in almost linear time. We illustrate the performances of our algorithm on a few synthetic and real data sets.

6.2.9. Analysis and Visualization of Maps Between Shapes
Participants: Frédéric Chazal, Maks Ovsjanikov.

In collaboration with L. Guibas (Stanford University), M. Ben Chen (Technion).

In this work we propose a method for analyzing and visualizing individual maps between shapes, or collections of such maps [23]. Our method is based on isolating and highlighting areas where the maps induce significant distortion of a given measure in a multi-scale way. Unlike the majority of prior work which focuses on discovering maps in the context of shape matching, our main focus is on evaluating, analyzing and visualizing a given map, and the distortion(s) it introduces, in an efficient and intuitive way. We are motivated primarily by the fact that most existing metrics for map evaluation are quadratic and expensive to compute in practice, and that current map visualization techniques are suitable primarily for global map understanding, and typically do not highlight areas where the map fails to meet certain quality criteria in a multi-scale way. We propose to address these challenges in a unified way by considering the functional representation of a map, and performing spectral analysis on this representation. In particular, we propose a simple multi-scale method for map evaluation and visualization, which provides detailed multi-scale information about the distortion induced by a map, which can be used alongside existing global visualization techniques.

6.2.10. Map-Based Exploration of Intrinsic Shape Differences and Variability
Participants: Frédéric Chazal, Maks Ovsjanikov.

In collaboration with L. Guibas and Raif Rustamov (Stanford University), M. Ben Chen and O. Azencot (Technion).

We develop a novel formulation for the notion of shape differences, aimed at providing detailed information about the location and nature of the differences or distortions between the two shapes being compared [27]. Our difference operator, derived from a shape map, is much more informative than just a scalar global shape similarity score, rendering it useful in a variety of applications where more refined shape comparisons are necessary. The approach is intrinsic and is based on a linear algebraic framework, allowing the use of many common linear algebra tools (e.g, SVD, PCA) for studying a matrix representation of the operator. Remarkably, the formulation allows us not only to localize shape differences on the shapes involved, but also to compare shape differences across pairs of shapes, and to analyze the variability in entire shape collections based on the differences between the shapes. Moreover, while we use a map or correspondence to define each shape difference, consistent correspondences between the shapes are not necessary for comparing shape differences, although they can be exploited if available. We give a number of applications of shape differences, including parameterizing the intrinsic variability in a shape collection, exploring shape collections using local variability at different scales, performing shape analogies, and aligning shape collections.

6.2.11. An operator Approach to Tangent Vector Field Processing
Participants: Frédéric Chazal, Maks Ovsjanikov.

In collaboration with M. Ben Chen and O. Azencot (Technion).

In this work [34], we introduce a novel coordinate-free method for manipulating and analyzing vector fields on discrete surfaces. Unlike the commonly used representations of a vector field as an assignment of vectors to the faces of the mesh, or as real values on edges, we argue that vector fields can also be naturally viewed as operators whose domain and range are functions defined on the mesh. Although this point of view is common in differential geometry it has so far not been adopted in geometry processing applications. We recall the theoretical properties of vector fields represented as operators, and show that composition of vector fields with other functional operators is natural in this setup. This leads to the characterization of vector field properties through commutativity with other operators such as the Laplace-Beltrami and symmetry operators, as well as to a straight-forward definition of differential properties such as the Lie derivative. Finally, we demonstrate a range of applications, such as Killing vector field design, symmetric vector field estimation and joint design on multiple surfaces.

6.3. Data Structures and Robust Geometric Computation

6.3.1. *The Stability of Delaunay triangulations*

Participants: Jean-Daniel Boissonnat, Ramsay Dyer.

In collaboration with Arijit Ghosh (Indian Statistical Institute)

We introduce a parametrized notion of genericity for Delaunay triangulations which, in particular, implies that the Delaunay simplices of δ-generic point sets are thick [45]. Equipped with this notion, we study the stability of Delaunay triangulations under perturbations of the metric and of the vertex positions. We quantify the magnitude of the perturbations under which the Delaunay triangulation remains unchanged.

6.3.2. *Delaunay Stability via Perturbations*

Participants: Jean-Daniel Boissonnat, Ramsay Dyer.

In collaboration with Arijit Ghosh (Indian Statistical Institute)

We present an algorithm that takes as input a finite point set in Euclidean space, and performs a perturbation that guarantees that the Delaunay triangulation of the resulting perturbed point set has quantifiable stability with respect to the metric and the point positions [43]. There is also a guarantee on the quality of the simplices: they cannot be too flat. The algorithm provides an alternative tool to the weighting or refinement methods to remove poorly shaped simplices in Delaunay triangulations of arbitrary dimension, but in addition it provides a guarantee of stability for the resulting triangulation.

6.3.3. *Deletions in 3D Delaunay Triangulation*

Participant: Olivier Devillers.

In collaboration with Kevin Buchin (Technical University Eindhoven, The Netherlands), Wolfgang Mulzer (Freie Universität Berlin, Germany), Okke Schrijvers, (Stanford University, USA) and Jonathan Shewchuk (University of California at Berkeley, USA)

Deleting a vertex in a Delaunay triangulation is much more difficult than inserting a new vertex because the information present in the triangulation before the deletion is difficult to exploit to speed up the computation of the new triangulation.

The removal of the tetrahedra incident to the deleted vertex creates a hole in the triangulation that need to be retriangulated. First we propose a technically sound framework to compute incrementally a triangulation of the hole vertices: *the conflict Delaunay triangulation*. The conflict Delaunay triangulation matches the hole boundary and avoid to compute extra tetrahedra outside the hole. Second, we propose a method that uses *guided randomized reinsertion* to speed up the point location during the computation of the conflict triangulation. The hole boundary is a polyhedron, this polyhedron is simplified by deleting its vertices one by one in a random order maintaining a polyhedron called *link Delaunay triangulation*, then the points are inserted in reverse order into the conflict Delaunay triangulation using the information from the link Delaunay triangulation to avoid point location [30].

6.3.4. A Convex Body with a Chaotic Random Polytope

Participants: Olivier Devillers, Marc Glisse, Rémy Thomasse.

Consider a sequence of points in a convex body in dimension d whose convex hull is dynamically maintained when the points are inserted one by one, the convex hull size may increase, decrease, or being constant when a new point is added. Studying the expected size of the convex hull when the points are evenly distributed in the convex is a classical problem of probabilistic geometry that yields to some surprising facts. For example, although it seems quite natural to think that the expected size of the convex hull is increasing with n the number of points, this fact is only formally proven for n big enough [16]. The asymptotic behavior of the expected size is known to be logarithmic for a polyhedral body and polynomial for a smooth one. If for a polyhedral or a smooth body, the asymptotic behavior is *somehow* "nice" it is possible to construct strange convex objects that have no such nice behaviors and we exhibit a convex body, such that the behavior of the expected size of a random polytope oscillates between the polyhedral and smooth behaviors when n increases [51].

6.3.5. Delaunay Triangulations and Cycles on Closed Hyperbolic surfaces

Participants: Mikhail Bogdanov, Monique Teillaud.

This work [40] is motivated by applications of *periodic* Delaunay triangulations in the Poincaré disk conformal model of the hyperbolic plane $\mathbb{H}^2$. A periodic triangulation is defined by an infinite point set that is the image of a finite point set by a (non commutative) discrete group G generated by hyperbolic translations, such that the hyperbolic area of a Dirichlet region is finite (i.e., a cocompact Fuchsian group acting on $\mathbb{H}^2$ without fixed points).

We consider the projection of such a Delaunay triangulation onto the closed orientable hyperbolic surface $M = \mathbb{H}^2/G$. The graph of its edges may have cycles of length one or two. We prove that there always exists a finite-sheeted covering space of M in which there is no cycle of length ≤ 2. We then focus on the group defining the Bolza surface (homeomorphic to a torus having two handles), and we explicitly construct a sequence of subgroups of finite index allowing us to exhibit a covering space of the Bolza surface in which, for any input point set, there is no cycle of length one, and another covering space in which there is no cycle of length two. We also exhibit a small point set such that the projection of the Delaunay triangulation on the Bolza surface for any superset has no cycle of length ≤ 2.

The work uses mathematical proofs, algorithmic constructions, and implementation.

6.3.6. Universal Point Sets for Planar Graph Drawings with Circular Arcs

Participant: Monique Teillaud.

In collaboration with Patrizio Angelini (Roma Tre University), David Eppstein (University of California, Irvine), Fabrizio Frati (The University of Sydney), Michael Kaufmann (MPI, Tübingen), Sylvain Lazard (EPI VEGAS), Tamara Mchedlidze (Karlsruhe Institute of Technology), and Alexander Wolff (Universität Würzburg).

We prove that there exists a set S of n points in the plane such that every n-vertex planar graph G admits a plane drawing in which every vertex of G is placed on a distinct point of S and every edge of G is drawn as a circular arc. [25]

6.3.7. A Generic Implementation of dD Combinatorial Maps in CGAL

Participant: Monique Teillaud.

In collaboration with Guillaume Damiand (Université de Lyon, LIRIS, UMR 5205 CNRS)

We present a generic implementation of dD combinatorial maps and linear cell complexes in CGAL. A combinatorial map describes an object subdivided into cells; a linear cell complex describes the linear geometry embedding of such a subdivision. In this paper [49], we show how generic programming and new techniques recently introduced in the C++11 standard allow a fully generic and customizable implementation of these two data structures, while maintaining optimal memory footprint and direct access to all information. To the best of our knowledge, the CGAL software packages presented here [59], [60] offer the only available generic implementation of combinatorial maps in any dimension.

6.3.8. *Silhouette of a Random Polytope*
Participant: Marc Glisse.

In collaboration with Sylvain Lazard and Marc Pouget (EPI VEGAS*) and Julien Michel (LMA-Poitiers).*

We consider random polytopes defined as the convex hull of a Poisson point process on a sphere in $\mathbb{R}^3$ such that its average number of points is n. We show [52] that the expectation over all such random polytopes of the maximum size of their silhouettes viewed from infinity is $\Theta(\sqrt{n})$.

6.3.9. *A New Approach to Output-Sensitive Voronoi Diagrams and Delaunay Triangulations*
Participant: Donald Sheehy.

In collaboration with Gary Miller (Carnegie Mellon University)

We describe [35] a new algorithm for computing the Voronoi diagram of a set of n points in constant-dimensional Euclidean space. The running time of our algorithm is $O(f \log n \log \Delta)$ where f is the output complexity of the Voronoi diagram and Δ is the spread of the input, the ratio of largest to smallest pairwise distances. Despite the simplicity of the algorithm and its analysis, it improves on the state of the art for all inputs with polynomial spread and near-linear output size. The key idea is to first build the Voronoi diagram of a superset of the input points using ideas from Voronoi refinement mesh generation. Then, the extra points are removed in a straightforward way that allows the total work to be bounded in terms of the output complexity, yielding the output sensitive bound. The removal only involves local flips and is inspired by kinetic data structures.

6.3.10. *A Fast Algorithm for Well-Spaced Points and Approximate Delaunay Graphs*
Participant: Donald Sheehy.

In collaboration with Gary Miller and Ameya Velingker (Carnegie Mellon University)

We present [32] a new algorithm that produces a well-spaced superset of points conforming to a given input set in any dimension with guaranteed optimal output size. We also provide an approximate Delaunay graph on the output points. Our algorithm runs in expected time $O(2^{O(d)}(n \log n + m))$, where n is the input size, m is the output point set size, and d is the ambient dimension. The constants only depend on the desired element quality bounds.

To gain this new efficiency, the algorithm approximately maintains the Voronoi diagram of the current set of points by storing a superset of the Delaunay neighbors of each point. By retaining quality of the Voronoi diagram and avoiding the storage of the full Voronoi diagram, a simple exponential dependence on d is obtained in the running time. Thus, if one only wants the approximate neighbors structure of a refined Delaunay mesh conforming to a set of input points, the algorithm will return a size $2^{O(d)}m$ graph in $2^{O(d)}(n \log n + m)$ expected time. If m is superlinear in n, then we can produce a hierarchically well-spaced superset of size $2^{O(d)}n$ in $2^{O(d)}n \log n$ expected time.

6.3.11. *Geometric Separators and the Parabolic Lift*
Participant: Donald Sheehy.

A geometric separator for a set U of n geometric objects (usually balls) is a small (sublinear in n) subset whose removal disconnects the intersection graph of U into roughly equal sized parts. These separators provide a natural way to do divide and conquer in geometric settings. A particularly nice geometric separator algorithm originally introduced by Miller and Thurston has three steps: compute a centerpoint in a space of one dimension higher than the input, compute a conformal transformation that "centers" the centerpoint, and finally, use the computed transformation to sample a sphere in the original space. The output separator is the subset of S intersecting this sphere. It is both simple and elegant. We show [36] that a change of perspective (literally) can make this algorithm even simpler by eliminating the entire middle step. By computing the centerpoint of the points lifted onto a paraboloid rather than using the stereographic map as in the original method, one can sample the desired sphere directly, without computing the conformal transformation.

<h1 style="text-align:center;color:red;">GRACE Project-Team</h1>

6. New Results

6.1. Diffusion layers for block ciphers

MDS matrices allow the construction of optimal linear diffusion layers in block ciphers. However, MDS matrices usually have a large description (for example, they can never be sparse), and this results in costly software/hardware implementations. We can solve this problem using *recursive MDS matrices*, which can be computed as a power of a simple companion matrix—and thus have a compact description suitable for constrained environments. Until now, finding recursive MDS matrices required an exhaustive search on families of companion matrices; this clearly limited the size of MDS matrices that one could look for. We have found a new direct construction, based on shortened BCH codes, which allows us to efficiently construct these matrices for arbitrary parameter sizes.

6.2. Rank metric codes over the rationals

Rank metric and Gabidulin codes over the rationals promise interesting applications to space-time coding. We have constructed optimal codes similar to Gabidulin codes but with complex coefficients, using number fields and Galois automorphisms. Using these codes we can completely bypass intermediate constructions using finite fields, which were the stumbling-block in classic constructions.

6.3. Cryptanalysis of McEliece cryptosystems based on Generalised Reed–Solomon codes

The McEliece encryption scheme based on binary Goppa codes was one of the first public-key encryption schemes [24]. Niederreiter [25] dramatically reduced the (huge) key size—a major problem with McEliece's original proposal—using Generalised Reed–Solomon (GRS) codes, but his modified scheme was broken by Sidelnikov and Shestakov [26] in 1992. There have been several attempts at repairing these smaller-key McEliece schemes. In collaboration with P. Gaborit, V. Gautier, A. Otmani and J.-P. Tillich, Alain Couvreur found polynomial time attacks on these schemes using the distinguishability of GRS codes from random codes.

6.4. New Identities relating Goppa codes

Goppa codes are strongly related to AG codes based on curves of genus 0. Among other applications, these codes are very famous for their cryptographic potential: they are one of the very few families of algebraic codes proposed for the McEliece encryption scheme which have not been broken up to now. At least for this reason, getting further knowledge on the structure of such codes is of interest. In [19], Alain Couvreur, A. Otmani and J.-P. Tillich proved a new identity yielding many improvements in the designed parameters of Goppa codes.

6.5. Root finding algorithms over local rings

Guillaume Quintin, in collaboration with J. Berthomieu and G. Lecerf, has developed new algorithms computing the roots of polynomials over complete local unramified rings [7]; this is important in the second stage of Guruswami–Sudan list decoding algorithms for codes over finite rings. Quintin has implemented these algorithms in MATHEMAGIX, using his FINITEFIELDZ and QUINTIX librairies.

6.6. Codes over rings

M. Barbier, C. Chabot and Guillaume Quintin proposed a new description for quasi–cyclic codes using the ring of matrices with polynomial entries, thus defining the new class of *quasi-BCH* codes. Guillaume Quintin proved that these codes can be regarded as interleaved subcodes of Reed–Solomon codes; this allowed them to define a polynomial-time decoding algorithm for quasi-BCH codes. Guillaume Quintin also generalized list decoding algorithms to codes over non commutative rings [8].

6.7. Quantum LDPC codes

For some time it was feared that quantum computers could not be built because of distortions of quantum states due to interaction with the environment. This issue could be addressed by the use of quantum codes. *Quantum LDPC codes* are very interesting candidates here, because their very fast decoding algorithm allows high error correction rates. But the design of good quantum LDPC codes is far more complicated than for their classical counterparts, and cannot be done by random generation. The best-known constructions come from algebraic topology and simplicial homology, but their limits were unknown. Nicolas Delfosse used Riemannian geometry theorems of Gromov to prove that an $[[n, k, d]]$-quantum code constructed from the homology of a simplicial surface satisfies $kd^2 \leq C(\log k)^2 n$ for some constant C [21].

Color codes are quantum LDPC codes constructed from 3–regular surface tilings whose set of faces is 3–colorable. Delfosse used morphisms of chain complexes to prove that the decoding of a color code can be reduced to the decoding of three associated surface codes; hence, every decoding algorithm for surface codes yields a decoding algorithm for color codes. From this result, Delfosse obtained theoretical lower bounds on the error threshold of a family of color codes [20].

6.8. New families of fast elliptic curves

Benjamin Smith has pioneered the use of mod-p reductions of Q-curves to produce elliptic curves with efficient scalar multiplication algorithms—which translates into faster encryption, decryption, signing, and signature verification operations on these curves. A theoretical article was presented at ASIACRYPT 2013 [9], and the Journal of Cryptology has invited the submission of a longer version. The theory was put into practice in collaboration with Craig Costello (Microsoft Research) and Huseyin Hisil (Yasar University). Their resulting publicly available implementation, which represents the state of the art in constant-time (side-channel conscious) elliptic curve scalar multiplication on 64-bit Intel platforms at the 128-bit security level, can carry out a constant-time scalar multiplication in 145k cycles on Ivy Bridge architectures. This work will appear in EUROCRYPT 2014 [17].

6.9. Tensor rank of multiplication over finite fields

Determining the tensor rank of multiplication over finite fields is a problem of great interest in algebraic complexity theory, but it also has practical importance: it allows us to obtain multiplication algorithms with a low bilinear complexity, which are of crucial significance in cryptography. In collaboration with S. Ballet and J. Chaumine [12], Julia Pieltant obtained new asymptotic bounds for the symmetric tensor rank of multiplication in finite extensions of finite fields $\mathbb{F}_q$. In the more general (not-necessarily-symmetric) case, Pieltant and H. Randriam obtained new uniform upper bounds for multiplication in extensions of $\mathbb{F}_q$. They also gave purely asymptotic bounds substantially improving those coming from uniform bounds, by using a family of Shimura curves defined over $\mathbb{F}_q$. This work will appear in Mathematics of Computation [22].

GRAND-LARGE Project-Team

5. New Results

5.1. Automated Code Generation for Lattice Quantum Chromodynamics

Participants: Denis Barthou, Konstantin Petrov, Olivier Brand-Foissac, Olivier Pène, Gilbert Grosdidier, Michael Kruse, Romain Dolbeau, Christine Eisenbeis, Claude Tadonki.

This ongoing work is about a Domain Specific Language which aims to simplify Monte-Carlo simulations and measurements in the domain of Lattice Quantum Chromodynamics. The tool-chain, called Qiral, is used to produce high-performance OpenMP C code from LaTeX sources. We discuss conceptual issues and details of implementation and optimization. The comparison of the performance of the generated code to the well-established simulation software is also made.[33][20][37]

5.2. A Fine-grained Approach for Power Consumption Analysis and Prediction

Participants: Alessandro Ferreira Leite, Claude Tadonki, Christine Eisenbeis, Alba Cristina de Melo.

Power consumption has became a critical concern in modern computing systems for various reasons including financial savings and environmental protection. With battery powered devices, we need to care about the available amount of energy since it is limited. For the case of supercomputers, as they imply a large aggregation of heavy CPU activities, we are exposed to a risk of overheating. As the design of current and future hardware is becoming more and more complex, energy prediction or estimation is as elusive as that of time performance. However, having a good prediction of power consumption is still an important request to the computer science community. Indeed, power consumption might become a common performance and cost metric in the near future. A good methodology for energy prediction could have a great impact on power-aware programming, compilation, or runtime monitoring. In this paper, we try to understand from measurements where and how power is consumed at the level of a computing node. We focus on a set of basic programming instructions, more precisely those related to CPU and memory. We propose an analytical prediction model based on the hypothesis that each basic instruction has an average energy cost that can be estimated on a given architecture through a series of micro-benchmarks. The considered energy cost per operation includes all of the overhead due to context of the loop where it is executed. Using these precalculated values, we derive an linear extrapolation model to predict the energy of a given algorithm expressed by means of atomic instructions. We then use three selected applications to check the accuracy of our prediction method by comparing our estimations with the corresponding measurements obtained using a multimeter. We show a 9.48% energy prediction on sorting.[35]

5.3. Switcheable scheduling

Participants: Lénaïc Bagnères, Cédric Bastoul, Taj Khan.

Parallel applications used to be executed alone until their termination on partitions of supercomputers. The recent shift to multicore architectures for desktop and embedded systems is raising the problem of the coexistence of several parallel programs. Operating systems already take into account the *affinity* mechanism to ensure a thread will run only onto a subset of available processors (e.g., to reuse data remaining in the cache since its previous execution). But this is not enough, as demonstrated by the large performance gaps between executions of a given parallel program on desktop computers running several processes. To support many parallel applications, advances must be made on the system side (scheduling policies, runtimes, memory management...). However, automatic optimization and parallelization can play a significant role by generating programs with dynamic-auto-tuning capabilities to adapt themselves to the complete execution context, including the system load.

Our approach is to design at compile-time programs that can adapt at run-time to the execution context. The originality of our solution is to rely on *switcheable scheduling*, a selected set of program restructuring which allows to swap between program versions at some meeting points without backtracking. A first step selects pertinent versions according to their performance behavior on some execution contexts. The second step builds the auto-adaptive program with the various versions. Then at runtime the program selects the best version by a low overhead sampling and profiling of the versions, ensuring every computation is useful.

This work has been started at Paris-Sud University by Cédric Bastoul before he joined the Inria CAMUS project team during this year. The first results have been presented in 2013 at the HiPEAC System Week and at the Rencontres Françaises de Compilation.

5.4. Solving Navier-Stokes equations on heterogeneous parallel architectures

Participants: Marc Baboulin, Jack Dongarra, Joël Falcou, Yann Fraigneau, Olivier Lemaître, Yushan Wang.

The Navier-Stokes equations describe a large class of fluid flows but are difficult to solve analytically because of their nonlinearity. We implemented a parallel solver for the 3-D Navier-Stokes equations of incompressible unsteady flows with constant coefficients, discretized by the finite difference method. We applied the prediction-projection method which transforms the Navier-Stokes equations into three Helmholtz equations and one Poisson equation. For each Helmholtz system, we applied the Alternating Direction Implicit (ADI) method resulting in three tridiagonal systems. The Poisson equation is solved using partial diagonalization which transforms the Laplacian operator into a tridiagonal one. Our implementation is based on MPI where the computations are performed on each subdomain and information is exchanged on the interfaces, and where the tridiagonal system solutions are accelerated using vectorization techniques. We provided performance results on a current multicore system.[31]

5.5. Optimizing NUMA effects in dense linear algebra software

Participants: Marc Baboulin, Adrien Rémy, Brigitte Rozoy, Masha Sosonkina.

We studied the impact of non-uniform memory accesses (NUMA) on the solution of dense general linear systems using an LU factorization algorithm. In particular we illustrated how an appropriate placement of the threads and memory on a NUMA architecture can improve the performance of the panel factorization and consequently accelerate the global LU factorization. We applied these placement strategies and presented performance results for a hybrid multicore/GPU LU algorithm as it is implemented in the public domain library MAGMA.

<p align="center" style="color:red">IN-SITU Project-Team</p>

6. New Results

6.1. Interaction Techniques

Participants: Caroline Appert, Michel Beaudouin-Lafon, David Bonnet, Anastasia Bezerianos, Olivier Chapuis [correspondant], Cédric Fleury, Stéphane Huot, Can Liu, Wendy Mackay, Halla Olafsdottir, Cyprien Pindat, Theophanis Tsandilas.

We explore interaction techniques in a variety of contexts, including individual interaction techniques on different display surfaces that range from mobile devices to very large wall-sized displays through standard desktop and tabletops. This year, we investigated how people can use different body parts and limbs to convey information to interactive systems. *BodyScape* provides a framework for analysing and designing interaction techniques that involve the entire human body. Both *WristPointing*, which overcomes the limited range of motion of the wrist, and *HeadPad*, which takes the user's head orientation into account, are whole body techniques that facilitate target acquisition. *Arpege* can interpret a wide range of chord gestures, designed according to the range of motion and limitations of the human hand, and includes a dynamic guide with integrated feedforward/feedback to enhance learning by novices, without slowing down experts. On mobile devices, we designed novel interaction techniques that increase the expressivity of gestures by a single finger, including *ThumbRock*, based on movement dynamics, *SidePress*, which senses pressure on the device, and *Powerup*, which detects proximity. We also continued to develop advanced interactive visualization techniques, including *Gimlenses*, which supports focus+context representations for navigating within 3D scenes.

BodyScape – The entire human body plays a central role in interaction. The BodyScape design space [34] (honorable mention at CHI 2013) explores the relationship between users and their environment, specifically how different body parts enhance or restrict movement for specific interactions. BodyScape can be used to analyze existing techniques or suggest new ones. In particular, we used it to design and compare two free-hand techniques, on-body touch and mid-air pointing, first separately, then combined. We found that touching the torso is faster than touching the lower legs, since it affects the user's balance; and touching targets on the dominant arm is slower than targets on the torso because the user must compensate for the applied force.

HeadPad – Rich interaction with high-resolution wall displays is not limited to remotely pointing at targets. Other relevant types of interaction include virtual navigation, text entry, and direct manipulation of control widgets. However, most techniques for remotely acquiring targets with high precision have studied remote pointing in isolation, focusing on pointing efficiency and ignoring the need to support these other types of interaction. We investigated high-precision pointing techniques capable of acquiring targets as small as 4 millimeters on a 5.5 meters wide display while leaving up to 93 of a typical tablet device's screen space available for task-specific widgets [27]. We compared these techniques to state-of-the-art distant pointing techniques and have shown that two of our techniques, a purely relative one and one that uses head orientation, perform as well or better than the best pointing-only input techniques while using a fraction of the interaction resources.

WristPointing – Wrist movements are physically constrained and take place within a small range around the hand's rest position. We explored pointing techniques that deal with the physical constraints of the wrist and extend the range of its input without making use of explicit mode-switching mechanisms [33]. Taking into account elastic properties of the human joints, we investigated designs based on rate control. In addition to pure rate control, we examine a hybrid technique that combines position and rate-control and a technique that applies non-uniform position-control mappings. Our experimental results suggest that rate control is particularly effective under low-precision input and long target distances. Hybrid and non-uniform position-control mappings, on the other hand, result in higher precision and become more effective as input precision increases.

Arpege – While multi-touch input has become a standard for interacting with devices equipped with a touchscreen with simple techniques like pinch-to-zoom, the number of gestures systems are able to interpret remains rather small. Arpège [23] is a progressive multitouch input technique for learning chords, as well as a robust recognizer and guidelines for building large chord vocabularies. We conducted two experiments to evaluate our approach. Experiment one validated our design guidelines and suggests implications for designing vocabularies, i.e. users prefer relaxed to tense chords, chords with fewer fingers and chords with fewer tense fingers. Experiment two demonstrated that users can learn and remember a large chord vocabulary with both Arpège and cheat sheets, and Arpège encourages the creation of effective mnemonics.

ThumbRock – Compared with mouse-based interaction on a desktop interface, touch-based interaction on a mobile device is quite limited: most applications only support tapping and dragging to perform simple gestures. Finger rolling provides an alternative to tapping but uses a recognition process that relies on either per-user calibration, explicit delimiters or extra hardware, making it difficult to integrate into current touch-based mobile devices. We introduce ThumbRock [19], a ready-to-use micro gesture that consists in rolling the thumb back and forth on the touchscreen. Our algorithm recognizes ThumbRocks with more than 96% accuracy without calibration nor explicit delimiter by analyzing the data provided by the touch screen with a low computational cost. The full trace of the gesture is analyzed incrementally to ensure compatibility with other events and to support real-time feedback. This also makes it possible to create a continuous control space as we illustrate with our MicroSlider, a 1D slider manipulated with thumb rolling gestures.

SidePress – Virtual navigation on a mobile touchscreen is usually performed using finger gestures: drag and flick to scroll or pan, pinch to zoom. While easy to learn and perform, these gestures cause significant occlusion of the display. They also require users to explicitly switch between navigation mode and edit mode to either change the viewport's position in the document, or manipulate the actual content displayed in that viewport, respectively. SidePress [31] augments mobile devices with two continuous pressure sensors co-located on one of their sides (Figure 9 -(Left)). It provides users with generic bidirectional navigation capabilities at different levels of granularity, all seamlessly integrated to act as an alternative to traditional navigation techniques, including scrollbars, drag-and-flick, or pinch-to-zoom. We built a functional hardware prototype and developed an interaction vocabulary for different applications. We conducted two laboratory studies. The first one showed that users can precisely and efficiently control SidePress; the second, that SidePress can be more efficient than drag-and-flick touch gestures when scrolling large documents.

Powerup – Current technology like Arduino (http://arduino.cc/) opens a large space for designing new electronic device. We built the Power-up Button [30] by combining both pressure and proximity sensing to enable gestural interaction with one thumb (Figure 9 -(Right)). Combined with a gesture recognizer that takes the hand's anatomy into account, the Power-up Button can recognize six different mid-air gestures performed on the side of a mobile device. This gives it, for instance, enough expressive power to provide full one-handed control of interface widgets displayed on screen. This technology can complement touch input, and can be particularly useful when interacting eyes-free. It also opens up a larger design space for widget organization on screen: the button enables a more compact layout of interface components than what touch input alone would allow. This can be useful when, e.g., filling the numerous fields of a long Web form, or for very small devices.

Gimlenses – Complex 3D virtual scenes such as CAD models of airplanes and representations of the human body are notoriously hard to visualize. Those models are made of many parts, pieces and layers of varying size, that partially occlude or even fully surround one another. Gimlenses [28] provides a multi-view, detail-in-context visualization technique that enables users to navigate complex 3D models by interactively drilling holes into their outer layers to reveal objects that are buried, possibly deep, into the scene (see Figure 10). These holes are constantly adjusted so as to guarantee the visibility of objects of interest from the parent view. Gimlenses can be cascaded and constrained with respect to one another, providing synchronized, complementary viewpoints on the scene. Gimlenses enable users to quickly identify elements of interest, get detailed views of those elements, relate them, and put them in a broader spatial context.

Dashboard Exploration – Visual stories help us communicate knowledge, share and interpret experiences and have become a focus in visualization research in recent years. In this paper we discuss the use of storytelling

../../../../projets/in-situ/IMG/sidepress.png ../projets/in-situ/IMG/powerup.png

Figure 9. (left) The SidePress prototype has two pressure sensors on one of its sides ; (right) The Power-Up button prototype is equipped with a button for sensing both pressure and proximity.

in Business Intelligence (BI) analysis [21] (Best Paper Award). We derive the actual practices in creating and sharing BI stories from in-depth interviews with expert BI analysts (both story "creators" and "readers"). These interviews revealed the need to extend current BI visual analysis applications to enable storytelling, as well as new requirements related to BI visual storytelling. Based on these requirements we designed and implemented a storytelling prototype tool with appropriate interaction techniques, that is integrated in an analysis tool used by our experts, and allows easy transition from analysis to story creation and sharing. We report experts' recommendations and reactions to the use of the prototype to create stories, as well as novices' reactions to reading these stories.

Hybrid-Image Visualizations – Data analysis scenarios often incorporate one or more displays with sufficiently large size and resolution to be comfortably viewed by different people from various distances. Hybrid-image visualizations [15] blend two different visual representations into a single static view, such that each representation can be perceived at a different viewing distance. They can thus be used to enhance overview tasks from a distance and detail-in-context tasks when standing close to the display. Viewers interact implicitly with these visualizations by walking around the space. By taking advantage of humans' perceptual capabilities, hybrid-image visualizations show different content to viewers depending on their placement, without requiring tracking of viewers in front of a display. Moreover, because hybrid-images use a perception-based blending approach, visualizations intended for different distances can each utilize the entire display.

Evolutionary Visual Exploration – In a high-dimensionality context, the visual exploration of information is challenging, as viewers are often faced with a large space of alternative views on the data. We present [14], a system that combines visual analytics with stochastic optimization to aid the exploration of multidimensional datasets characterized by a large number of possible views or projections. Starting from dimensions whose values are automatically calculated by a PCA, an interactive evolutionary algorithm progressively builds (or evolves) non-trivial viewpoints in the form of linear and non-linear dimension combinations, to help users discover new interesting views and relationships in their data. The system calibrates a fitness function (optimized by the evolutionary algorithm) to take into account the user interactions to calculate and propose new views. Our method leverages automatic tools to detect interesting visual features and human interpretation to derive meaning, validate the findings and guide the exploration without having to grasp advanced statistical concepts. Our prototype was evaluated through an observational study with five domain experts, and helped

../../../../projets/in-situ/IMG/gimlenses-teaser.png

Figure 10. Exploring the CAD drawing of a car engine. The three Gimlenses provide detailed views of different constituent parts of the engine, at different magnification levels and with varying orientation, while revealing their location inside the global 3D model. (a) Context view. (b) Magnified side view of a knot behind, and thus originally hidden by, the cylinder head cover. (c) View fully revealing a poppet valve in-context from a different angle than the main view, with (d) another Gimlens configured so as to provide a low-angled point of view on the valve.

them quantify qualitative hypotheses, try out different scenarios to dynamically transform their data, and to better formulate their research questions and build new hypotheses for further investigation.

6.2. Research Methods

Participants: Michel Beaudouin-Lafon, Anastasia Bezerianos, Jérémie Garcia, Stéphane Huot, Ilaria Liccardi, Wendy Mackay [correspondant].

Conducting empirical research is a fundamental part of InSitu's research activities, including observation of users in field and laboratory settings to discover problems faced by users, controlled laboratory experiments to evaluate the effectiveness of the technologies we develop, longitudinal field studies to determine how our technologies work in the real world, and participatory design, to explore design possibilities with users throughout the design process.

However, we not only use research methods, we also investigate and develop them. As organizers of the CHI'13 conference in Paris, which had record-breaking numbers of submissions (over 2000) and participants (3500), we instituted a number of innovations in both the process of creating the program and presenting information to conference participants. In collaboration with researchers at MIT, we introduced an "author-sourcing" process (with an 87% participation rate) for collecting affinity data. We then developed a collaborative, interactive, visualization system on the WILD wall display, combined with the *Cobi* interactive constraint-solving system, that enabled us to resolve all presenter conflicts and successfully place all 500+ papers and events in relevant sessions ([35], [26]). We also replaced the "CHI Madness" series of 25-second presentations with "Video Previews", in which each research paper, course, panel or other event has a 30-second video preview. These are now available on the CHI'13 website, the ACM/CHI YouTube channel and in the ACM Digital Library, before the paywall. We also developed and field-tested the Interactive Schedule on large, interactive displays, which allowed conference attendees to both view upcoming Video Previews and use their mobile phones to search for particular content and create customized playlists [29]. We also developed two interactive table-top applications that were presented at CHI'13 Interactivity, that allowed attendees to visualize and explore conference events as well as to create customized video playlists.

The RepliCHI workshop at CHI'13, co-organized by Wendy Mackay, examined issues with respect to encouraging replication of controlled experiments, and introduced the RepliCHI award to top research articles that offer strong empirical contributions that include replication. She also organized a session called *Interacting with CHI* in which participants explained the technologies and processes they developed to support the CHI conference design and execution.

In the context of our work with Interactive Paper to support music composition, we developed Paper Tonnetz, a paper-based interface to composing melodies and chords based on musical patterns expressed in Euler's Tonnetz, and demonstrated it at CHI'13 Interactivity ([22]). We also explored how to create an interactive event for the "Fête de la Science", called "Design Me a Sound Landscape", in which participants can create their own ways of expressing a landscape and add natural sounds, such as wind, rain, moving water, that another participant can experience as they move on an interactive floor. Finally, we explored the drawing process, with the *Drawing Assistant* ([25]) in which users receive guidance and feedback as they learn to draw from photographs.

6.3. Engineering of interactive systems

Participants: Caroline Appert, Michel Beaudouin-Lafon [correspondant], Olivier Chapuis, Stéphane Huot, Wendy Mackay.

InSitu has a long tradition of developing software tools and user interface toolkits to facilitate the creation of interactive systems. These tools allow us to better experiment with our ideas and are therefore an integral part of our research methodology. Most of them are freely available and some are used outside InSitu for research or teaching.

Our work has focused on developing middleware for the WILD platform, InSitu's experimental ultra-high-resolution interactive room for studying collaborative interaction and the visualization of very large datasets [2]. WILD features a wall-sized display with 32 monitors, a multitouch table, a motion-tracking system and various mobile devices. Running applications on WILD requires developing advanced distributed systems that coordinate, in real time, the 16 computers of the cluster driving the wall display with a variety of clients and servers running on other computers, including mobile devices.

We investigated the use of Web standards and protocols to develop and deploy such applications. Hydrascope [24] introduces the concept of *meta-application* that combines, adapts and/or repurposes existing web applications for an environment such as WILD. It uses a web browser (or even a web engine, e.g. WebKit) as a rendering and interaction toolkit and Web protocols (HTTP and WebSockets) for communication. We demonstrated how to control a wall-size presentation tool built on Google Present and a wall-size map built on Google Maps without modifying these applications but by taking advantage of the capability of web applications for introspection.

This approach was used to develop CHIWall, an application designed to help us schedule the CHI 2013 conference that InSitu chaired this year in Paris. The resulting tool combines a wall-size display of the full program with a constraint-detection and constraint-solving assistant called Cobi [26], which itself uses crowd-sourced information from the authors. The resulting application supports collaborative work to fine-tune the program (Figure 11) and features a flexible architecture that has been reused for other prototype applications.

In summary, InSitu has continued to make significant progress towards mature toolkits that support post-WIMP and distributed user interfaces. These toolkits, in turn, have enabled us to experiment with novel interaction techniques using rapid prototyping. Conversely, our work on novel interaction techniques has driven the development of software toolkits that embody their underlying principles, facilitating further exploration. This back-and-forth between techniques, methods and tools is a defining feature of InSitu, captured by the *Designeering Interaction* [11] framework. As the focus of our research on interaction techniques has shifted from on-the-desktop to off-the-desktop, this approach has proven more valid than ever: improving interaction in such environments requires more complex software architectures and tools; in turn, these tools and architectures are a key step to getting these technologies outside the lab.

../../../../projets/in-situ/IMG/CHIwall.png

Figure 11. A group of users interacting with the CHIwall application to fine tune the schedule of the CHI 2013 conference (500 events (400 research articles) in 200 sessions, in 16 parallel tracks over four days). The tablet interface provides additional detail about the content of a session, such as the abstract and the details of the affinity data crowd-sourced from the authors.

<p align="center" style="color:red">**M3DISIM Team**</p>

6. New Results

6.1. Modeling

6.1.1. Collective effect in molecular motors assembly

Participants: Matthieu Caruel, Jean-Marc Allain [Ecole Polytechnique], Lev Truskinovsky [Ecole Polytechnique].

Skeletal muscles consist of active material capable of producing force. At the microscale, force is the result of complex interactions between two types of proteins named actin and myosin that work coherently in very large assemblies ($\sim 10^9$). The passive mechanical response of skeletal muscles at fast time scales is dominated by long range interactions inducing cooperative behavior without breaking the detailed balance. This leads to such unusual "material properties" as negative equilibrium stiffness and different behavior in force and displacement controlled loading conditions. Our fitting of experimental data suggests that "muscle material" is finely tuned to perform close to a critical point which explains large fluctuations observed in muscles close to the stall force. See paper [28].

6.1.2. Dimensional reductions of a cardiac model for effective validation and calibration

Participants: Matthieu Caruel, Alexandre Imperiale, Radomir Chabiniok [King's College London], Philippe Moireau, Dominique Chapelle, Yves Lecarpentier [Institut du Cœur].

Complex 3D beating heart models are now available, but their complexity makes calibration and validation very difficult tasks. We thus propose a systematic approach of deriving simplified reduced-dimensional models, in "0D"— typically, to represent a cardiac cavity, or several coupled cavities and in "1D"—to model elongated structures such as muscle samples or myocytes. We apply this approach with an earlier-proposed 3D cardiac model designed to capture length-dependence effects in contraction, which we here complement by an additional modeling component devised to represent length-dependent relaxation. We then present experimental data produced with rat papillary muscle samples when varying preload and afterload conditions, and we achieve some detailed validations of the 1D model with these data, including for the length-dependence effects that are accurately captured. Finally, when running simulations of the 0D model pre-calibrated with the 1D model parameters, we obtain pressure–volume indicators of the left ventricle in good agreement with some important features of cardiac physiology, including the so-called Frank–Starling mechanism, the End-Systolic Pressure–Volume Relationship, as well as varying elastance properties. This integrated multi-dimensional modeling approach thus sheds new light on the relations between the phenomena observed at different scales and at the local versus organ levels. See papers [13], [22].

6.1.2.1. Surface-based electrophysiology modeling and assessment of physiological simulations in atria

Participants: Dominique Chapelle, Annabelle Collin, Jean-Frédéric Gerbeau [Reo Project-Team], Mélèze Hocini [Institut LIRYC - IHU Bordeaux], Michel Haïssaguerre [Institut LIRYC - IHU Bordeaux].

The objective of this work is to assess a previously-proposed surface-based electrophysiology model with detailed atrial simulations. This model – derived and substantiated by mathematical arguments – is specifically designed to address thin structures such as atria, and to take into account strong anisotropy effects related to fiber directions with possibly rapid variations across the wall thickness. The simulation results are in excellent adequacy with previous studies, and confirm the importance of anisotropy effects and variations thereof, see Figure 1 . Furthermore, this surface-based model provides dramatic computational benefits over 3D models with preserved accuracy. See paper [23].

6.1.3. Strong convergence results in the asymptotic behavior of the 3D-shell model

Participants: Dominique Chapelle, Annabelle Collin.

Figure 1. Surface-based modeling of atrial electrophysiology

The objective of this work is to revisit the asymptotic convergence properties – with respect to the thickness parameter – of the earlier-proposed 3D-shell model. This shell model is very attractive for engineering applications, in particular due to the possibility of directly using a general 3D constitutive law in the corresponding finite element formulations. We establish strong convergence results for the 3D-shell model in the two main types of asymptotic regimes, namely, bending- and membrane-dominated behavior. This is an important achievement, as it completely substantiates the asymptotic consistency of the 3D-shell model with 3D linearized isotropic elasticity. See paper [14].

6.1.4. *Mechanical modeling and numerical methods for poromechanics: Applications to cardiac perfusion*

Participants: Bruno Burtschell, Dominique Chapelle, Philippe Moireau.

We have previously formulated a rather general modeling framework of poromechanics – formulations that combine solid and fluid components to represent the behavior of a porous medium – to take into account large deformations and rapid fluid flows, see [29]. This allows to consider, in particular, the application of blood perfusion within the cardiac tissue, which – indeed – features these specific complex phenomena, out of the scope of classical poromechanical models. One of our major objectives now, within the PhD of Bruno Burtschell, is to propose and assess some associated relevant numerical schemes, which requires special care regarding both space and time discretizations. In a first stage, in order to ease numerical prototyping and assessments, an axisymmetric reduction of the model has been formulated, and some existing algorithms of fluid-structure interaction have been implemented within this axisymmetric framework (in FreeFEM++). The rationale is that our poromechanics formulations feature some rather deep similarities to so-called Arbitrary-Lagrangian-Eulerian (ALE) fluid-structure formulations, hence the latter are considered as a natural starting point for further extensions.

6.1.5. *Personalized modeling for cardiac amyloidosis diagnosis*

Participants: Alessandro Felder, Dominique Chapelle, Philippe Moireau, Jean-François Deux [Hôpital Henri Mondor], Thibault Damy [Hôpital Henri Mondor].

Cardiac amyloidosis is a condition induced by pathological deposition of amyloid proteins within muscle tissue and nerves, thus severely impairing the cardiac function and often requiring cardiac transplantation as the only available treatment. Our objective here in a first stage is to use our previously developed patient-specific modeling methodologies to analyse some clinical cases – based on actual patient data – to better apprehend the impact of the pathology on biomechanical properties. Further perspectives include the modeling of the protein deposition and associated tissue remodeling in order to predict the disease evolution in a patient-specific context. This work is performed in collaboration with medical doctors from Hôpital Henri Mondor (Créteil).

6.2. Model-Data Interaction

6.2.1. *State observers of a vascular fluid-structure interaction model through measurements in the solid*

Participants: Cristobal Bertoglio [Reo Project-Team], Dominique Chapelle, Miguel Angel Fernández [Reo Project-Team], Jean-Frédéric Gerbeau [Reo Project-Team], Philippe Moireau.

We analyze the performances of two types of Luenberger observers – namely, the so-called Direct Velocity Feedback and Schur Displacement Feedback procedures, originally devised for elasto-dynamics – to estimate the state of a fluid-structure interaction model for hemodynamics, when the measurements are assumed to be restricted to displacements or velocities in the solid. We first assess the observers using hemodynamics-inspired test problems with the complete model, including the Navier-Stokes equations in Arbitrary Lagrangian-Eulerian formulation, in particular. Then, in order to obtain more detailed insight we consider several well-chosen simplified models, each of which allowing a thorough analysis – emphasizing spectral considerations – while illustrating a major phenomenon of interest for the observer performance, namely, the added mass effect for the structure, the coupling with a lumped-parameter boundary condition model for the fluid flow, and the fluid dynamics effect *per se*. Whereas improvements can be sought when additional measurements are available in the fluid domain in order to more effectively deal with strong uncertainties in the fluid state, in the present framework this establishes Luenberger observer methods as very attractive strategies – compared, e.g., to classical variational techniques – to perform state estimation, and more generally for uncertainty estimation since other observer procedures can be conveniently combined to estimate uncertain parameters. See paper [11].

6.2.2. *Improving Efficiency of Data Assimilation Procedure for a Biomechanical Heart Model by Representing Surfaces as Currents*

Participants: Alexandre Imperiale, Alexandre Routier [Aramis Team], Philippe Moireau, Stanley Durrleman [Aramis Team].

We adapt the formalism of currents to compare data surfaces and surfaces of a mechanical model and we use this discrepancy measure to feed a data assimilation procedure. We apply our methodology to perform parameter estimation in a biomechanical model of the heart using synthetic observations of the endo- and epicardium surfaces of an infarcted left ventricle. We compare this formalism with a more classical signed distance operator between surfaces and we numerically show that we have improved the efficiency of our estimation justifying the use of state-of-the-art computational geometry formalism in the data assimilation measurements processing. See paper [24].

6.2.3. *Optimal observer for parabolic problems*

Participants: Karine Mauffrey, Philippe Moireau.

Figure 2. Heart model immersed in MR image data

We aim at proposing optimal observers strategies for reconstructing the solution of general systems of PDEs using avalaible observations, including both wave-type equations and heat-like equations. The main objective of this work is to present a self-contained analysis. For a general parabolic system, we have established the exponential stability of the operator occurring in the equation satisfied by the error between the target and the optimal observer. The proof relies on two major hypotheses: an observability inequality satisfied by the observation operator and a controllability property for the modeling error operator by which model noises enter the dynamics (controllability property which is related to the invertibility of the solution of the associated infinite dimensional Riccati equation). The next questions we want to tackle are the discretisation of the model and the construction of a reduced Kalman filter.

<p style="text-align:center">**Maxplus Project-Team**</p>

6. New Results

6.1. Théorie spectrale max-plus et géométrie métrique/Max-plus spectral theory and metric geometry

6.1.1. Introduction

Participants: Marianne Akian, Stéphane Gaubert, Cormac Walsh.

Étant donné un noyau $a : S \times S \to \mathbb{R} \cup \{-\infty\}$, on peut lui associer le problème spectral max-plus

$$\sup_{y \in S} a(x,y) + u(y) = \lambda + u(x), \quad \forall x \in S, \tag{1}$$

dans lequel on cherche le vecteur propre $u : S \to \mathbb{R} \cup \{-\infty\}$ et la valeur propre correspondante $\lambda \in \mathbb{R} \cup \{-\infty\}$. Comme nous l'avons rappelé dans les §3.2 et 3.3 , le problème spectral (9) intervient en contrôle ergodique: l'ensemble S est l'espace des états, et l'application $a(x,y)$ fournit le gain associé à la transition $x \to y$. Le cas où S est fini est classique, l'on a alors un résultat précis de représentation de l'espace propre, à l'aide d'un certain graphe, dit graphe critique. Des résultats existent également lorsque S est compact et que le noyau vérifie certaines propriétés de régularité.

Dans [61], nous avons considéré le cas où S est non compact. Lorsque $\lambda = 0$, l'espace propre est analogue à l'espace des fonctions harmoniques défini en théorie (classique ou probabiliste) du potentiel. En introduisant l'analogue max-plus de la frontière de Martin, nous avons obtenu un analogue de la formule de représentation de Poisson des fonctions harmoniques : toute solution u de (9) peut être représentée sous la forme :

$$u = \sup_{w \in \mathcal{M}_m} w + \mu_u(w) , \tag{2}$$

où $\mathcal{M}_m \subset (\mathbb{R} \cup \{-\infty\})^S$ est l'analogue max-plus de la frontière de Martin minimale (l'ensemble des fonctions harmoniques extrémales normalisées), et où μ_u joue le rôle de la mesure spectrale. Nous avons montré aussi que les éléments de l'espace de Martin minimal peuvent être caractérisés comme les limites de "quasi-géodésiques". La frontière de Martin max-plus généralise dans une certaine mesure la frontière d'un espace métrique construite à partir des horo-fonctions (fonctions de Busemann généralisées), ou horo-frontière. Ces résultats inspirent les travaux des sections suivantes, qui portent sur des cas remarquables d'espaces métriques (§6.1.3) ou sur des applications en théorie des jeux (§6.1.2).

<p style="text-align:center">***English version***</p>

Let the kernel $a : S \times S \to \mathbb{R} \cup \{-\infty\}$ be given. One may associate the max-plus spectral equation (9), where the eigenvector $u : S \to \mathbb{R} \cup \{-\infty\}$ and the eigenvalue $\lambda \in \mathbb{R} \cup \{-\infty\}$ are unknown. As we recalled in §3.2 and refmonotone, this spectral problem arises in ergodic optimal control: the set S is the *state space*, and the map $a(x,y)$ is the *transition reward*. The case when S is finite is classical, a precise spectral theorem is known, with a characterisation of the eigenspace in terms of a critical graph. Some results have been shown when S is compact, assuming that the kernel a satisfies some regularity properties.

In [61], we considered the case where S is non-compact. When $\lambda = 0$, the eigenspace is analoguous to the set of harmonic functions defined in classical or probabilistic potential theory. By introducing a max-plus analogue of the classical Martin boundary, we obtained an analogue of the Poisson representation of harmonic functions, showing that any solution u of (9) may be represented as in (10) where $\mathcal{M}_m \subset (\mathbb{R} \cup \{-\infty\})^S$ is a max-plus analogue of the minimal Martin boundary (the set of normalised extremal harmonic functions), and μ_u plays the role of the spectral measure. We also showed that the elements of the minimal Martin boundary can be characterised as limits of certain "almost-geodesics". The max-plus Martin boundary generalises to some extent the boundary of metric spaces defined in terms of horofunctions (generalised Busemann functions), or horoboundary. These results have inspired the work of the next sections, which deal either with interesting examples of metric spaces (§6.1.3) or with applications to zero-sum games (§6.1.2).

6.1.2. *Asymptotiques d'itérées d'applications contractantes au sens large et jeux à somme nulle en horizon long/Asymptotics of iterates of nonexpansive mappings and zero-sum games*

Participants: Jérôme Bolte, Stéphane Gaubert, Guillaume Vigeral.

On s'intéresse ici à l'existence du paiement moyen pour les jeux répétés, et plus généralement, à l'existence du vecteur de "taux de fuite" $\lim_k f^k(x)/k$ où f est une application de $\mathbb{R}^n$ dans lui même, nonexpansive pour une norme quelconque. Dans le cas particulier des jeux, f est un opérateur de Shapley, qui est nonexpansif pour la norme sup. On montre dans [45] que la limite existe si l'application f est définissable dans une structure o-minimale. Ceci généralise des résultats de Bewley, Kohlberg, et Neyman, qui montraient que la limite existe si f est semi-algébrique. L'extension au cas o-minimal permet notamment de traiter des opérateurs de type "log-exp" apparaissant en contrôle sensible au risque. Ce travail traite aussi de la question de savoir si un jeu dont les fonctions de paiement et de transition sont définissables dans une structure o-minimale admet un opérateur de Shapley f définissable. Un contre exemple montre que f n'est pas forcément définissable dans la même structure, mais l'on montre qu'il en est ainsi dès que les probabilités de transition ont une structure séparable.

English version

We study the question of the existence of the mean payoff for repeated games, and more generally, the existence of a vector of "escape rates", $\lim_k f^k(x)/k$, where f is a self-map of $\mathbb{R}^n$, non-expansive in some norm. In the special case of zero-sum games, f is a Shapley operator, and it is sup-norm nonexpansive. We showed in [45] that this limit does exist as soon as the map f is definable in an o-minimal structure. This generalizes results of Bewley, Kohlberg, and Neyman, who showed that this limit exists if f is semi-algebraic. The extension to the case of o-minimal structures allows one in particular to deal with log-exp type operators arising in risk sensitive control. This work also adresses the question of knowing whether a game with definable payment and transition functions has a Shapley operator that is definable in the same structure. We gave a counter example showing that this may not be the case, but showed that the Shapley operator is definable as soon as the transition probabilities have a separable structure.

6.1.3. *Isométries de la géométrie de Hilbert/Isometries of the Hilbert geometry*

Participants: Cormac Walsh, Bas Lemmens [Kent University, UK].

L'un des intérêts de l'horo-frontière est de renseigner sur le groupe des isométries d'un espace métrique. En effet, ce groupe agit naturellement sur l'horo-frontière, et cette action peut parfois être mieux comprise que l'action du groupe sur l'espace d'origine.

Nous avons utilisé ces idées pour étudier le groupe des isométries pour la métrique de Hilbert. De La Harpe [179] a donné plusieurs conjectures relatives à ce groupe. Nous avons montré dans [51], en utilisant l'horo-frontière, que le groupe des isométries est exactement le groupe des transformations linéaires projectives à moins que le domaine ne soit une coupe d'un cône symétrique non-Lorentzien. Dans ce dernier cas, le groupe linéaire projectif est d'index 2 dans le groupe des isométries. Le cas particulier où le domaine est un polytope a été traité précédemment dans [136].

Dans [51] nous déterminons aussi le groupe des isométries pour une métrique fortement reliée à la métrique de Hilbert, à savoir la métrique de Thompson sur un cone.

English version

One use for the horofunction boundary is to study the group of isometries of a metric space. This is because this group has a well defined action on the horoboundary and it is likely that in many cases this action will be easier to understand than the action on the space itself.

We have been applying these ideas to investigate the isometries of the Hilbert geometry. De La Harpe [179] has previously made several conjectures about the isometry group of this space. We have shown [51] using the horofunction boundary that the isometry group is exactly the group of projective linear transformations unless the domain on which the geometry is defined is a cross section of a non-Lorentzian symmetric cone, in which case the projective linear group is of index two in the isometry group.

The special case when the domain is a polytope was previously considered in [136].

In the paper [51], we also determine the isometry group of closely related metric, the Thompson geometry on a cone.

6.1.4. *Consensus non-commutatif et contraction d'opérateurs de Kraus/Noncommutative consensus and contraction of Kraus maps*

Participants: Stéphane Gaubert, Zheng Qu.

Dans le travail [47], on s'est intéressé à la vitesse de convergence vers l'équilibre d'une itération de la forme $x^{k+1} = T(x^k)$, $x^k \in X$, où T est une application linéaire préservant un cône dans un espace de Banach X, telle que $T(e) = e$, pour un certain vecteur e dans l'interieur du cône. On s'intéresse aussi à l'itération dans l'espace dual, $y^{k+1} = T^*(y^k)$, $y^k \in X^*$, lorsque $\langle y^0, e \rangle = 1$.

Le cas classique est celui où $T(x) = Px$ est un opérateur de Markov. L'itération primale traduit alors la convergence vers le "consensus", et l'itération duale traduit la convergence de la distribution de probabilité en temps k vers l'état stationnaire. Dans ce cas, le taux de contraction (en un coup) $\kappa(P)$ d'une itération primale, pour la semi-norme de Hilbert $\|z\|_H := \max_i z_i - \min_j z_j$, ainsi que le taux de contraction d'une itération duale, pour la métrique en variation totale, coïncident et sont caractérisés par une formule dûe à Doeblin et Dobrushin (coefficient d'ergodicité),

$$\kappa(P) := 1 - \min_{i,j} \sum_{s=1}^n \min(P_{is}, P_{js}).$$

On a donné ici une généralisation de cette formule au cas d'opérateurs abstraits, qui s'applique en particulier aux opérateurs de Kraus qui interviennent en information quantique. Ces derniers opérent sur l'espace des matrices symétriques, et sont de la forme

$$T(x) = \sum_k a_k x a_k^* \qquad \text{avec} \qquad \sum_k a_k a_k^* = I \ .$$

Dans [34], nous avons étendu ces résultats aux flots non-linéaires sur les cones.

English version

In a recent work [27], we studied the speed of convergence to equilibrium of an iteration of the form $x^{k+1} = T(x^k)$, $x^k \in X$, where T is a linear map preserving a cone in a Banach space X, such that $T(e) = e$, for some vector e in the interior of the cone. We also considered the iteration in the dual space X^*, $y^{k+1} = T^*(y^k)$, $y^k \in X^*$, where $\langle y^0, e \rangle = 1$.

The classical application arises when $T(x) = Px$ is a Markov operator. Then, the primal iteration represents the dynamics of consensus, whereas the dual iteration represents the evolution of the probability distribution as a function of time. Then, the (one-shot) contraction rate $\kappa(P)$ of the primal iteration, with respect to Hilbert's seminorm $\|z\|_H := \max_i z_i - \min_j z_j$, and the contraction rate of the dual iteration, with respect to the total variation metric, coincide, and are characterized by a formula of Doeblin and Dobrushin (ergodicity coefficient),

$$\kappa(P) := 1 - \min_{i,j} \sum_{s=1}^n \min\left(P_{is}, P_{js}\right).$$

We gave here a generalization of this formula to an abstract operators on a cone. This covers in particular the Kraus maps arising in quantum information theory. The latter maps act on the space of symmetric matrices. They can be written as

$$T(x) = \sum_k a_k x a_k^* \qquad \text{with} \qquad \sum_k a_k a_k^* = I \ .$$

In [34], we generalized these results to non-linear flows over cones.

6.2. Algèbre linéaire max-plus et convexité abstraite/Max-plus linear algebra and abstract convex analysis

6.2.1. *Convexité max-plus ou tropicale/Max-plus or tropical convexity*

Participants: Xavier Allamigeon, Stéphane Gaubert, Eric Goubault [CEA], Ricardo Katz [Conicet, Argentine].

On étudie les analogues max-plus ou tropicaux des ensembles convexes. Ceux-ci sont utiles en particulier pour représenter de manière effective les ensembles d'états accessibles de systèmes à événements discrets [9], ils sont aussi apparus récemment en géométrie tropicale, dans toute une série de travaux à la suite de Sturmfels et Develin [96]. Les polyèdres max-plus peuvent aussi être vus comme des limites de déformations de polyèdres classiques, sur lesquels ils donnent un éclairage de nature combinatoire. Toutes ces motivations ont inspiré la recherche d'analogues des résultats fondamentaux d'analyse convexe classique: séparation, projection, points extrémaux, à la suite en particulier de [8].

Dans un travail de X. Allamigeon, S. Gaubert, et E. Goubault [64], [16], on a mis en évidence un critère combinatoire pour la caractérisation des sommets des polyèdres tropicalement convexes. Celui-ci s'exprime à l'aide d'hypergraphes orientés, et de leurs composantes fortement connexes. Ce critère possède la propriété d'être vérifiable en un temps presque linéaire en la taille de l'hypergraphe.

On en déduit un analogue tropical de la méthode de la double description [16] (méthode très utilisée sur les polyèdres classiques, et dûe à Motzkin *et al.* [148]). Cet algorithme permet de calculer les sommets d'un polyèdre défini de façon externe (intersection de demi-espaces ou d'hyperplans tropicaux). Grâce au critère combinatoire précédent, l'algorithme améliore de plusieurs ordres de grandeur les techniques connues jusqu'alors. Ceci est confirmé par de nombreuses expérimentations. Ce travail est motivé par des applications à l'analyse statique [63] et aux systèmes à événements discrets [99], dans lesquelles la manipulation de tels polyèdres est le goulot d'étranglement.

Il est connu qu'un polyèdre tropical peut être représenté comme l'enveloppe convexe d'un ensemble minimal de points et rayons, donnés par ses sommets et ses rayons extrêmes [111]. Dans un travail réalisé par X. Allamigeon et R. Katz [17], et effectué en partie lors de visites de R. Katz à Inria, on étudie la question duale de la caractérisation des représentations minimales par demi-espaces. On montre qu'un polyèdre tropical possède *essentiellement* une unique représentation minimale par demi-espaces, lorsque leurs apex appartiennent au polyèdre. On montre que les apex de ces demi-espaces non-redondants correspondent à certains sommets du complexe tropical introduit par Develin et Sturmfels [96]. On introduit également un critère combinatoire pour l'élimination de demi-espaces redondants à l'aide d'hypergraphes orientés.

Dans un travail en cours de X. Allamigeon, P. Benchimol, S. Gaubert et R. Katz, nous étudions la tropicalisation des représentations par demi-espaces des polyèdres convexes sur le corps des séries de Puiseux. Nous démontrons ainsi une conjecture de Develin et Yu [97]. Celle-ci assure qu'étant donné un polytope tropical pur, il existe un polytope *relevé* sur les séries de Puiseux, dont les demi-espaces associés aux faces se "tropicalisent" en une représentation par demi-espaces du polytope tropical initial.

Des applications de ces travaux à l'algorithmique, concernant en particulier les jeux répétés, sont discutées dans la Section 6.4.2 . Une application aux systèmes temps réel est discutée dans la Section 6.5.4 .

English version

We study the max-plus or tropical analogues of convex sets. These have been used in particular to represent effectively the accessible sets of certain discrete event systems [9]. They also appeared in tropical geometry, following the work of Sturmfels and Develin [96]. Max-plus polyhedra can be thought of as limits of deformations of classical polyhedra, on which they give a combinatorial insight. These motivations have inspired the investigation of analogues of basic results of classical convex analysis: separation, projection, representation by extreme points, following [8].

In a work of X. Allamigeon, S. Gaubert, and E. Goubault [16], we introduce a combinatorial criterion for the characterization of the vertices of tropically convex polyhedra. It is expressed in terms of directed hypergraphs and their strongly connected components. This criterion can be verified in almost linear time in the size of the hypergraph.

This allows to develop a tropical analogue of the double description method [16] (this method is widely used for classical convex polyhedra, and is due to Motzkin *et al.* [148]). This algorithm is able to determine all the vertices of a polyhedron defined externally (intersection of tropical half-spaces of hyperplanes). Thanks to the combinatorial criterion mentioned above, the algorithm improves the existing methods by several orders of magnitude. This is confirmed by several experiments. This is motivated by applications to static analysis [63] and discrete event systems [99], in which computing such polyhedra turns out to be the bottleneck.

It is well-known that a tropical polyhedron can be represented as the convex hull of a minimal set of points and rays, provided by its vertices and extreme rays [111]. In a work of X. Allamigeon and R. Katz [17], partly done during visits of R. Katz at Inria, the dual problem of characterizing the minimal representations by half-spaces is studied. We show that a tropical polyhedron admits *essentially* a unique minimal external representation by half-spaces, provided that their apices belong to the polyhedron. We prove that the apices of these half-spaces correspond to certain vertices of the tropical complex introduced by Develin and Sturmfels [96]. We also establish a combinatorial criterion allowing to eliminate redundant half-spaces using directed hypergraphs.

In an ongoing work of X. Allamigeon, P. Benchimol, S. Gaubert and R. Katz, we study the tropicalization of the representation by half-spaces of convex polyhedra over the field of Puiseux series. In particular, we prove a conjecture of Develin and Yu [97]. It states that, given a pure tropical polytope, there exists a lifting polytope over Puiseux series, such that the facet-defining half-spaces are "tropicalized" into a representation by half-spaces of the initial polytope.

Some algorithmic applications of this work concerning in particular mean payoff games, will be discussed in Section 6.4.2 . Applications to real time systems will be discussed in Section 6.5.4 .

6.2.2. *Systèmes linéaires max-plus/Max-plus linear systems*

Participants: Marianne Akian, Stéphane Gaubert, Alexander Guterman [Moscow State University].

Dans [37], on montre des formules de Cramer pour des systèmes linéaires sur diverses extensions du semi-anneau max-plus. Les éléments de ces extensions sont des nombres tropicaux enrichis d'une information de multiplicité, de signe ou d'angle par exemple. On obtient ainsi des résultats d'existence et d'unicité qui généralisent plusieurs résultats de [121], [153], [107], [161], [127]. De plus, pour certaines extensions du semi-anneau max-plus, les preuves fournissent des algorithmes de type Jacobi ou Gauss-Seidel pour résoudre les systèmes linéaires.

English version

In [37], we prove general Cramer type theorems for linear systems over various extensions of the tropical semiring, in which tropical numbers are enriched with an information of multiplicity, sign, or argument. We obtain existence or uniqueness results, which extend or refine earlier results in [121], [153], [107], [161], [127]. Moreover, some of our proofs lead to Jacobi and Gauss-Seidel type algorithms to solve linear systems in suitably extended tropical semirings.

6.3. Algèbre max-plus, déformations et asymptotiques /Max-plus algebra, deformations and asymptotic analysis

6.3.1. Introduction

Comme indiqué dans le §3.7 , l'algèbre max-plus est la limite d'une déformation de l'algèbre classique, ou plutôt du semi-corps des réels positifs. Elle peut aussi fournir des estimations de ces déformations, puisque

$$\max(a, b) \leq \epsilon \log\left(e^{a/\epsilon} + e^{b/\epsilon}\right) \leq \epsilon \log(2) + \max(a, b) \ . \tag{3}$$

L'utilisation de ces propriétés a déjà conduit dans le passé aux travaux sur les perturbations de valeurs propres [55], [54], [53], ou sur les grandes déviations [1], [59]. Dans les travaux qui suivent, nous exploitons ces propriétés dans des contextes reliés ou similaires à ceux de nos travaux précédents.

English version

As detailled in §3.7 , max-plus algebra is the limit of a deformation of classical algebra, or more precisely of the semi-field of usual real positive numbers. It can also give estimations for these deformations using for instance (11). By using these properties, we already obtained some works on singular perturbations of matrix eigenvalues [55], [54], [53], or on large deviations [1], [59]. In the works described below, we are exploiting again these properties in contexts that are related or similar to those of our earlier works.

6.3.2. Aspects tropicaux des algorithmes de scaling matriciel/Tropical aspects of matrix scaling problems

Participants: Marianne Akian, Stéphane Gaubert, Meisam Sharify Najafabadi [Univ. Manchester].

Une partie du travail de thèse de M. Sharify [167] portait sur les méthodes de mise à l'échelle pour améliorer la précision du calcul de valeurs propres. En appliquant les techniques de [53], [54], on montrait notamment que l'ordre de grandeur des valeurs propres d'un faisceau matriciel est donné (sous des conditions de non-dégénerescence) par les valeurs propres tropicales, qui peuvent être calculées de manière robuste, et fournissent ainsi une mise à l'échelle pour calculer les valeurs propres classiques.

Nous avons poursuivi ce travail dans [41]. On calcule cette fois l'ordre de grandeur des valeurs propres d'un polynôme matriciel au moyen des racines tropicales du polynôme obtenu en appliquant une norme donnée aux coefficients. Les racines dépendent de la norme choisie, et la norme de Frobenius est optimale en un certain sens. On obtient des bornes générales pour les ratios entre modules des valeurs propres et racines tropicales qui généralisent les bornes obtenues par Polya et Ostrowski dans le cas de polynômes scalaires. On raffine aussi ces bornes, en particulier lorsque les racines tropicales sont bien séparées les unes des autres.

English version

A part of the PhD work of M. Sharify [167] dealt with scaling methods to improve the accuracy of eigenvalue numerical computions. Applying the techniques of [53], [54], we showed in particular that the order of magnitude of the eigenvalues of a matrix pencil can be determined (under nondegeracy conditions) by computing tropical eigenvalues. The latter can always be computed accurately and provide a scaling which can be combined with standard numerical methods for matrix pencils.

We have pursued this work in [41]. Now, we compute the order of magnitude of the eigenvalues of a matrix polynomial by using the tropical roots of a polynomial obtained by applying a norm to the coefficients of the original matrix polynomial. The tropical roots depend on the chosen norm, and the Frobenius turns out to be optimal in a certain sense. We obtain indeed general bounds on the ratios between the modulus of the eigenvalues of the matrix polynomial and the tropical roots which generalize the bounds of Polya and Ostrowski available for scalar polynomials. We also improve these bounds, in particular when the tropical roots are well separated.

6.3.3. *Méthodes tropicales de localisation de valeurs propres de matrices/Tropical methods for the localisation of matrix eigenvalues*

Participants: Marianne Akian, Stéphane Gaubert, Andrea Marchesini.

Le travail de stage de M2 d'Andrea Marchesini a conduit à la publication [14] dans laquelle on montre des inégalités de type majorisation entre les valeurs propres d'une matrice et les valeurs propres tropicales de la matrice de ses modules. En particulier, la majoration est une généralisation de l'inégalité de Friedland [106] concernant le rayon spectral.

La thèse d'Andrea Marchesini s'inscrit dans le prolongement de son stage de M2 dans l'équipe et certains des travaux de la thèse de Meisam Sharify [167]. Le but est d'obtenir des inégalités de type majorisation permettant d'estimer a priori les valeurs propres de matrices ou de faisceaux de matrices, en faisant éventuellement intervenir des hypothèses de bon conditionnements. En particulier on recherche la localisation de ces valeurs propres en fonction de valeurs propres de matrices agrégées ou simplifiées. On cherchera aussi à obtenir le même type de localisation ou d'estimation dans le cas des vecteurs propres associés, par exemple en utilisant les techniques de compléments de Schur de [54] ou les idées de Murota [149].

L'idée est ensuite d'utiliser ces résultats de localisation pour améliorer la précision des algorithmes de calcul numérique de valeurs propres de matrices, en particulier en construisant des changements d'échelle exploitant les calculs tropicaux, à effectuer préalablement à l'appel d'algorithmes classiques comme QZ. Les travaux de Stéphane Gaubert et Meisam Sharify [115] ont montré l'intérêt de cette approche, notamment pour les problèmes de faisceaux quadratiques de valeurs propres issus de systèmes mécaniques pour lesquels on dispose de nombreux exemples pathologiques pour les algorithmes existants. Dans un travail en collaboration avec Françoise Tisseur et James Hook de l'Université de Manchester, on montre l'intérêt des changements d'échelle en termes de le conditionnement des valeurs propres.

English version

The M2 internship of Andrea Marchesini led to the publication [14], in which we show majorization type inequalities between the eigenvalues of a matrix and the tropical eigenvalues of the matrix obtained by applying the modulus entrywise. In particular, the bound is a generalization of the inequality of Friedland [106] concerning the spectral radius.

The PhD thesis follows his M2 internship and some of the works of Meisam Sharify's PhD thesis [167]. The aim is to obtain majorization type inequalities allowing one to estimate the eigenvalues of matrices or matrix polynomials, using possibly assumptions on condition numbers. In particular, one may look for estimates of these eigenvalues using the eigenvalues of aggregated or simplified matrices. One may also try to find the same type of estimates for the associated eigenvectors, for instance by using techniques of Schur complements from [54] or ideas of Murota [149].

One would like to use these estimation results to improve the accuracy of eigenvalue numerical computions, in particular by constructing scaling methods using tropical techniques, which may be used before calling usual algorithms as QZ. The works of Stéphane Gaubert and Meisam Sharify [115] showed the interest of this approach, in particular for quadratic matrix polynomials issued from mechanical systems for which there exists several pathological examples for existing algorithms. In a work with Françoise Tisseur and James Hook from Manchester University, we show the interest of these scaling methods on the eigenvalue conditioning.

6.3.4. *Mesures et applications maxitives/Maxitive measures and maps*

Participants: Marianne Akian, Stéphane Gaubert, Paul Poncet.

La thèse de Paul Poncet [154] concernait essentiellement ce que l'on appelle l'analyse idempotente, c'est-à dire l'étude des espaces fonctionnels ou linéaires de dimension infinie sur l'algèbre tropicale, ou tout autre semi-anneau idempotent. Paul Poncet a développé pour cela un point de vue treillis continus comme dans [1], ou plus généralement domaines. Depuis la soutenance, plusieurs articles issus du manuscrit de thèse sont en cours de publication ou de soumission, et d'autres travaux pousuivant ceux de la thèse sont en cours avec les membres de l'équipe.

La première partie de la thèse traitait des mesures maxitives, en particulier de l'existence d'une densité cardinale ou d'une densité d'une mesure par rapport à une autre (théorème de Radon-Nikodym), et de la régularité d'une mesure maxitive. Ces travaux sont publiés ou en cours de publication dans [49] et [23] respectivement.

La deuxième partie concernait les convexes dans les semi-treillis ou l'algèbre max-plus, pour lesquels Paul Poncet a pu établir des théorèmes de type Krein-Milman, réciproque de Milman, et représentation de Choquet. [48] traite du cas des semi-treillis.

On sait que les résultats sur les convexes tropicaux de dimension infinie de [154] permettent de retrouver partiellement les résultats sur la frontière de Martin max-plus décrits dans la section 6.1.1 . Dans un travail commun nous essayons d'obtenir d'autres applications et extensions du théorème de représentation de Choquet tropical. En particulier on considère le cas d'ensembles ordonnés qui ne sont pas forcément des treillis tels que le cône des matrices symmetriques positives muni de l'ordre de Loewner.

English version

The PhD thesis work of Paul Poncet [154] concerned essentially what is called idempotent analysis, that is the study of infinite dimensional functional or linear spaces over tropical algebra, or any other idempotent semiring. For this aim, Paul Poncet developed the point of view of continuous lattices, as in [1], or more generally of domains. Since the defense of his thesis, several papers derived from the thesis manuscript have been submitted and some are published or up to be published. Some other works pursuing the thesis work are done with team members.

The first part of the Paul Poncet's thesis concerned maxitive measures, in particular the existence of a cardinal density of a measure, or that of a density of a measure with respect to another (Radon-Nikodym theorem), and the regularity of a maxitive measure. These works are now published or accepted for publication in [49] and [23] respectively.

A second part concerned convex sets in lattices or max-plus algebra, for which Paul Poncet showed results such as a Krein-Milman type theorem, a Milman converse type theorem, and a Choquet representation type theorem. [48] concerns the case of semilattices.

We know that the results on infinite dimensional tropical convex sets of [154] allow one to recover at least partially the results on max-plus Martin boundaries described in Section 6.1.1 . In a joint work, we try to obtain other applications and extensions of the max-plus Choquet representation theorem. In particular, we consider the case of ordered sets that are not necessarily semilattices, such as the cone of nonnegative symmetric matrices endowed with the Loewner order.

6.4. Algorithmes/Algorithms

6.4.1. *Itération sur les politiques pour le contrôle stochastique et les jeux répétés à somme nulle/Policy iterations for stochastic control and repeated zero sum games*
Participants: Marianne Akian, Jean Cochet-Terrasson [CGA], Sylvie Detournay, Stéphane Gaubert.

L'algorithme d'itération sur les politiques est bien connu pour résoudre efficacement les équations de la programmation dynamique associées à des problèmes de contrôle stochastique avec critère à horizon infini (Howard) ou ergodique (Howard, et Denardo et Fox). Récemment, il a été généralisé au cas de problèmes de jeux à deux joueurs et somme nulle dégénérés (avec paiements ergodiques et de type "multi-chaîne"), au moyen de techniques d'algèbre max-plus et de théorie du potentiel non linéaire [84]. Chaque itération de base

de cet algorithme utilise la résolution d'un système d'équations linéaires dont l'opérateur est monotone, mais dont la taille peut être grande, soit parce qu'il provient d'une discrétisation fine d'une équation aux dérivées partielles, soit parce qu'il est associé à un problème discret de grande taille comme le graphe du Web.

La thèse de Sylvie Detournay [95] a permis de développer et d'étudier un algorithme associant une méthode d'itération sur les politiques du type de celle introduite par Cochet-Terrasson et Gaubert dans [84] et une méthode multigrille algèbrique, afin de résoudre des problèmes de jeux à somme nulle dégénérés, éventuellement posés directement sous forme discrète. L'ensemble des codes nouveaux associés, écrits en C, est déposé sur le projet "pigames" de la gforge et sera disponible librement.

Sylvie Detournay a en particulier implémenté et raffiné l'algorithme proposé dans [84], en l'associant soit à des méthodes de résolution exacte de systèmes linéaires, soit à des méthodes multigrilles algébriques, en utilisant aussi des méthodes multigrilles multiplicatives pour le calcul de la mesure invariante de chaînes de Markov irréductibles, comme celles introduites par De Sterck. Ceci a permis l'obtention de résultats numériques dans le cas de discrétisations d'équations d'Isaacs associées à des jeux de poursuite déterministes ou aléatoires. Cela a aussi permis de tester de manière systématique l'algorithme sur des instances aléatoires de jeux de type Richman. Certains de ces résultats, ainsi que la présentation de l'algorithme (de manière plus concrète que dans [84], et avec les détails d'implémentation) sont présentés dans [24]. Des details supplémentaires ainsi que la preuve de convergence de l'algorithme peuvent être trouvés dans [56].

Des résultats récents de Ye ainsi que Hansen, Miltersen et Zwick montrent que l'algorithme d'itération sur les politiques, restreint à la classe des jeux à somme nulle (à 1 ou 2 joueurs) actualisés de facteur d'actualisation donné, est fortement polynomial. Dans [40], [29], on montre que ceci est le cas aussi pour l'algorithme d'itération sur les politiques pour les jeux à somme nulle et paiement moyen, restreint à la classe des jeux qui ont temps moyen de retour ou d'arrivée à un état donné borné. La preuve utilise des techniques de théorie de Perron-Frobenius non-linéaire, permettant de ramener le problème à paiement moyen à un problème actualisé (de facteur d'actualisation dépendant de l'état et des actions). La même technique permet aussi de traiter le cas de jeux à somme nulle actualisés dont le facteur d'actualisation peut dépendre de l'état et des actions et prendre éventuellement des valeurs supérieures à 1.

English version

Policy iteration is a powerful and well known algorithm to solve the dynamic programming equation associated to stochatic control (one player game) problems with infinite horizon criterion (Howard) or ergodic criterion (Howard and Denardo and Fox). It has recently been extended to degenerate two players problems (with ergodic payoff and in "multichain" cases) using ideas from max-plus algebra and nonlinear potential theory [84]. One basic iteration of the algorithm consists in solving a linear system the operator of which is monotone, but with a size which may be large since it comes from the discretization of a partial differential equation or since it is associated to a large size discrete problem arising from instance from the Web graph.

The PhD thesis of Sylvie Detournay [95] developed and studied an algorithm for degenerate two player games (that may come from a discrete time and finite state space model) combining a policy iteration such as the one introduced in [84] by Cochet-Terrasson et Gaubert, and an algebraic multigrid method (AMG). All new corresponding algorithms, coded in C, belong to the gforge project "pigames" and will be distributed openly.

In particular, Sylvie Detournay has implemented and refined the algorithm proposed in [84], while associating it either to direct linear solvers, or to the AMG methods already used in the nondegenerate case, and using also multiplicative AMG methods for computing invariant measures of Markov chains, such as the one introduced by De Sterck. This allowed her to obtain numerical results in the case of discretisations of Isaacs equations associated to deterministic or stochastic pursuit games. This also allowed her to test systematically the algorithm on random instances of Richman type games.

Some of these results, together with the presentation of the algorithm (in a more practical manner than in [84], with implementation details) are gathered in [24]. Additional details and the convergence proof of the algorithm can be found in [56].

Recent results of Ye and Hansen, Miltersen and Zwick show that policy iteration for one or two player (perfect information) zero-sum stochastic games, restricted to instances with a fixed discount rate, is strongly polynomial. In [40], [29], we show that policy iteration for mean-payoff zero-sum stochastic games is also strongly polynomial when restricted to instances with bounded first mean return time to a given state. The proof is based on methods of nonlinear Perron-Frobenius theory, allowing us to reduce the mean-payoff problem to a discounted problem with state dependent discount rate. Our analysis also shows that policy iteration remains strongly polynomial for discounted problems in which the discount rate can be state dependent (and even negative) at certain states, provided that the spectral radii of the nonnegative matrices associated to all strategies are bounded from above by a fixed constant strictly less than 1.

6.4.2. *Algorithmique des polyèdres tropicaux/Algorithmics of tropical polyhedra*
Participants: Xavier Allamigeon, Pascal Benchimol, Stéphane Gaubert, Eric Goubault [CEA], Michael Joswig [TU Darmstadt].

X. Allamigeon, S. Gaubert, et E. Goubault, ont développé dans [63], [16] plusieurs algorithmes permettant de manipuler des polyèdres tropicaux. Ceux-ci correspondent aux travaux décrits dans §6.2.1 . Ils permettent notamment de déterminer les sommets et rayons extrêmes d'un polyèdre tropical défini comme intersection de demi-espaces, ou inversement, de calculer une représentation externe à partir d'un ensemble de générateurs. Ces algorithmes sont implémentés la bibliothèque TPLib (voir §5.3).

Dans un travail en cours de X. Allamigeon, P. Benchimol, S. Gaubert et M. Joswig, nous avons défini un analogue tropical de l'algorithme du simplexe qui permet de résoudre les problèmes de *programmation linéaire tropicale*, *i.e.*

$$
\begin{aligned}
\text{minimiser} \quad & \max_{1 \le j \le n} c_j + x_j \\
\text{sous les contraintes} \quad & \max\big(\max_{1 \le j \le n}(a_{ij}^+ + x_j), b_i^+\big) \ge \max\big(\max_{1 \le j \le n}(a_{ij}^- + x_j), b_i^-\big), \quad i = 1, \cdots, m \qquad (4) \\
& x \in (\mathbb{R} \cup \{-\infty\})^n
\end{aligned}
$$

où les entrées du programme $a_{ij}^\pm$, $b_i^\pm$, c_j sont à valeur dans $\mathbb{R} \cup \{-\infty\}$. Ces problèmes sont intimement liés à la résolution de jeux répétés à somme nulle, puisque résoudre un jeux à paiement moyen déterministe est équivalent à déterminer si un problème de programmation linéaire admet un point réalisable [57].

Comme son homologue usuel, le simplexe tropical pivote entre des points de base (tropicaux), jusqu'à atteindre l'optimum du programme linéaire. La différence fondamentale avec l'algorithme du simplexe classique est que le pivotage est réalisé de manière purement combinatoire, en s'appuyant sur des descriptions locales du polyèdre tropical défini par les contraintes à l'aide d'(hyper)graphes orientés. Ceci nous a permis de prouver que *l'étape de pivotage (incluant le calcul des coûts réduits) a la même complexité en temps que dans l'algorithme classique, i.e.* $O(n(m + n))$. Ceci est d'autant plus inattendu que la structure des arêtes tropicales entre deux points de base sont géométriquement plus complexes (elles sont constituées de plusieurs segments de droite, jusqu'à n).

Le simplexe tropical a la propriété d'être fortement corrélé avec l'algorithme du simplexe classique. Grâce au principe de Tarski, le simplexe usuel peut être transposé tel quel sur des programmes linéaires dont les coefficients en entrée sont non plus des réels, mais sur le corps $\mathbb{R}\{\{t\}\}$ des séries de Puiseux généralisées en une certaine indéterminée t, *i.e.* des objets de la forme :

$$
c_{\alpha_1} t^{\alpha_1} + c_{\alpha_2} t^{\alpha_2} + \cdots \qquad (5)
$$

où les α_i sont des réels, les coefficients c_{α_i} sont des réels non-nuls, et où la séquence des $\alpha_1, \alpha_2, \cdots$ est strictement croissante et soit finie, soit non-bornée. L'opposé du plus petit exposant de la série, $-\alpha_1$, est appelé *valuation* de la série. Un programme linéaire tropical est dit *relevé* en un problème linéaire sur $\mathbb{R}\{\{t\}\}$, si la valuation des coefficients en entrée de ce dernier sont égaux aux coefficients du problème tropical. Dans nos travaux, nous avons établi la correspondance suivante entre le simplexe usuel et le simplexe tropical : *pour tout programme linéaire tropical générique, l'algorithme du simplexe tropical trace l'image par la valuation du chemin sur l'algorithme du simplexe usuel sur n'importe quel relèvement du programme tropical dans $\mathbb{R}\{\{t\}\}$.*

Les résultats présentés ci-dessus sont rassemblés dans le preprint [43]. Ils ont fait l'objet de plusieurs présentations en conférence [32], [33].

Ces résultats ouvrent la possibilité de relier la complexité de l'algorithme du simplexe usuel avec celles des jeux déterministes. Pour ces derniers, on sait seulement que leur résolution est dans la classe de complexité NP ∩ coNP, et on ignore s'il existe un algorithme de complexité polynomiale. De façon similaire, on ne sait pas caractériser de façon précise la complexité de l'algorithme du simplexe usuel. Celle-ci dépend fortement de la règle de pivotage utilisée, et il existe des problèmes sur lesquelles de nombreuses règles de pivotage ont une complexité exponentielle. L'existence d'une règle de pivotage qui permettrait au simplexe de terminer en temps polynomial sur n'importe quelle instance est encore aujourd'hui une question ouverte.

Dans un deuxième travail, nous avons relié les deux problèmes ouverts précédents, grâce à l'algorithme du simplexe tropical. Nous avons en effet exhibé une classe de règles de pivotage, dites *combinatoires*, et avons montré qu'elles satisfont la propriété suivante : *s'il existe une règle de pivotage combinatoire qui permet de résoudre tout problème de programmation linéaire usuel en temps polynomial, alors on peut résoudre les jeux à paiement moyen en temps (fortement) polynomial.* Le terme *combinatoire* fait référence au fait que la règle est définie en fonction du signe des mineurs de la matrice des coefficients du problème linéaire.

Ce dernier résultat est décrit dans le preprint [42].

English version

X. Allamigeon, S. Gaubert, and E. Goubault, have developed in [63], [16] algorithms allowing one to manipulate tropical polyhedra. They correspond to the contributions described in §6.2.1 . In particular, they can be used to determine the vertices and extreme rays of a tropical polyhedron defined as the intersection of half-spaces, or inversely, to compute an external description from a set of generators. These algorithms are implemented in the library TPLib (see §5.3).

In an ongoing work of X. Allamigeon, P. Benchimol, S. Gaubert and M. Joswig, we introduced a tropical analogue of the simplex algorithm, allowing one to solve problems of *tropical linear programming*, which are of the form (12), where the coefficients of the program, $a_{ij}^{\pm}$, $b_i^{\pm}$, c_j take their values in the max-plus semiring $\mathbb{R} \cup \{-\infty\}$. These problems are closely related to mean payoff games, as solving a game of this kind is equivalent to determine whether a tropical linear program admits a feasible point [57].

Like the classical simplex algorithm, the tropical simplex algorithm performs pivoting operations between basis points, until it reaches the optimum. The main discrepancy with the classical algorithm is that the pivoting is now a purely combinatorial operation, which is performed by using a local description of the polyhedron by a directed hypergraph. This allowed us to show that *a tropical pivoting step (including computing reduced costs) has the same complexity as in the classical simplex algorithm, i.e. $O(n(m + n))$. This is all the more surprising as the tropical edge between two given points has a geometrically more complex structure in the tropical case (it is constituted of up to n ordinary line segments).

The tropical simplex algorithm turns out to be closely related to the classical one. Thanks to Tarski's principle, the latter is also valid for linear programs over the field $\mathbb{R}\{\{t\}\}$ of generalized Puiseux series in an indeterminate t. These series are of the form (13), where the α_i are real numbers, the coefficients c_{α_i} are non-zero reals, and the sequence $\alpha_1, \alpha_2, \cdots$ is strictly increasing and either finite or unbounded. The opposite of the smallest exponent of the series, $-\alpha_1$, is called *valuation*. A tropical linear program is said to be *lifted* to a linear program over $\mathbb{R}\{\{t\}\}$ if the valuation of the coefficients of the latter are sent to the coefficients of

the former by the valuation. We showed the following relation between the classical simplex algorithm and its tropical analogue: *for all generic tropical linear program, the tropical simplex algorithm computes the image by the valuation of the path of the classical simplex algorithm, applied to any lift in $\mathbb{R}\{\{t\}\}$ of the original program*.

These results are gathered in the preprint [43]. They have been presented in several conferences [32], [33].

They allow one to relate the complexity of the classical simplex algorithm with the complexity of mean payoff games. The latter is unsettled, these games are known to be in the class NP ∩ coNP but it is not known whether they can be solved in polynomial time. Basic complexity issues regarding the classical simplex algorithm are also unsettled: its execution time depends on the pivoting rule, and many pivoting rules have been shown to have exponential worst case behaviors. The existence of a pivoting rule leading the simplex to terminate in polynomial time is still an open question.

In a second work, we related these two open questions, via the tropical simplex algorithm. We identified a class of pivoting rules, which are said to be *combinatorial*, and show that they have the following property: *if there is a combinatorial pivoting rule allowing one to solve every classical linear programming problem in polynomial time, then, mean payoff games can be solved in (strongly) polynomial time.* By *combinatorial*, we mean that the rule depends only of the coefficients of the system through the signs of minors of the coefficients matrix.

This result is given in the preprint [42].

6.4.3. *Problèmes d'accessibilité dans les hypergraphes orientés et leur complexité/Reachability problems in directed hypergraphs and their complexity*

Participant: Xavier Allamigeon.

Les hypergraphes orientés sont une généralisation des graphes orientés, dans lesquelles chaque arc relie un ensemble de sommets à un autre. Ils jouent un rôle important dans les travaux récents sur la convexité tropicale (voir §6.2.1), puisqu'ils offrent une représentation naturelle des cônes définis sur le sous-semi-anneau booléen $\mathbb{B} = \{-\infty, 0\}$.

Dans un travail de X. Allamigeon [15], on étudie la complexité de problèmes d'accessibilité sur les hypergraphes orientés. Nous introduisons un algorithme de complexité presque linéaire permettant de déterminer les composantes fortement connexes terminales (qui n'accèdent à aucune autre composante si ce n'est elles-mêmes) d'un hypergraphe.

Nous établissons également une borne inférieure sur-linéaire sur la taille de la réduction transitive de la relation d'accessibilité dans les hypergraphes. Cela indique que la relation d'accessibilité dans les hypergraphes orientés est combinatoirement plus complexe que celle des graphes orientés. Cela suggère aussi que des problèmes comme le calcul des composantes fortement connexes est plus difficile sur les hypergraphes que sur les graphes. Nous mettons d'ailleurs en évidence une réduction en temps linéaire du problème du calcul des ensembles minimaux dans une famille d'ensembles donnée, vers le problème du calcul de toutes les composantes fortement connexes d'un hypergraphe. Le problème du calcul des ensembles minimaux a été largement étudié dans la littérature [155], [175], [174], [156], [157], [158], [101], [69], et aucune algorithme en temps linéaire n'est connu à ce jour.

English version

Directed hypergraphs are a generalization of directed graphs, in which the tail and the head of the arcs are sets of vertices. It appears that they play an important role in the recent works on tropical convexity (see §6.2.1), since they offer a natural representation of cones defined over the boolean sub-semiring $\mathbb{B} = \{-\infty, 0\}$.

In a work of X. Allamigeon [15], we study the complexity of reachability problems on directed hypergraphs. We introduce an almost linear-time algorithm allowing to determine the terminal strongly connected components (a component is said to be *terminal* when no other component is reachable from it).

We also establish a super-linear lower bound over the size of the transitive reduction of the reachability relation in directed hypergraphs. This indicates that the reachability relation is combinatorially more complex in directed hypergraphs than in directed graphs. This also suggests that reachability problems such as computing all strongly connected components are likely to be harder in hypergraphs than in graphs. Besides, we show that the minimal set problem can be reduced in linear time to the problem of computing all strongly connected components in hypergraphs. The former problem consists in finding all minimal sets among a given family of sets. It has been well studied in the literature [155], [175], [174], [156], [157], [158], [101], [69], and no linear time algorithm is known.

6.4.4. *Approximation max-plus de fonctions valeurs et équations de Riccati généralisées/Max-plus approximation of value functions and generalized Riccati equations*

Participants: Stéphane Gaubert, Zheng Qu, Shanjian Tang [Fudan University, Shanghai].

La thèse de Zheng Qu, supervisée par S. Gaubert et S. Tang, a porté sur le développement de méthodes tropicales en programmation dynamique approchée [12].

Les méthodes d'approximation max-plus conduisent à approcher la fonction valeur d'un problème de contrôle ou de jeux par un supremum d'un nombre fini de formes quadratiques, voir notamment [113]. On s'intéresse ici à l'analyse théorique (complexité) ainsi qu'à l'amélioration de ces méthodes. Dans certains cas, ces formes quadratiques sont propagées par des flots d'équations de Riccati généralisées. Afin d'effectuer des analyses d'erreur, on exploite les propriétés de contraction du flot de Riccati pour certaines métriques connues sur le cône des matrices positives, et en particulier pour la métrique de Thompson. Celle-ci n'est rien d'autre que $d_T(A, B) = \| \log \operatorname{spec}(A^{-1}B) \|_\infty$, où spec désigne la suite des valeurs propres d'une matrice, et $\log$ s'entend composante par composante.

Ceci nous a amené à étudier le problème général du calcul du taux de contraction d'un flot monotone sur un cône, pour la métrique de Thompson. En effet, les propriétés de contraction de l'équation de Riccati standard sont connues (résultats de Bougerol pour la métrique Riemanienne invariante, et de Wojtowski pour la métrique de Thompson), mais les techniques de preuve employées dans ce cadre (semigroupes de matrices symplectiques) ne s'étendent pas aux équations généralisées.

On donne dans [114], [28] une formule explicite générale pour le taux de contraction pour la métrique de Thompson d'un flot monotone, faisant seulement intervenir le générateur du flot et sa dérivée. On a notamment appliqué ce résultat à une équation de Riccati généralisée associé à des problèmes de contrôle stochastique avec critère quadratique, dans lesquels la dynamique comporte un terme bilinéaire en le contrôle et le bruit. On a montré dans ce cas que la métrique de Thompson est la seule métrique de Finsler invariante pour laquelle le flot est nonexpansif, et l'on a caractérisé la constante de contraction locale.

Une application de ces résultats à l'analyse d'une méthode de réduction de la malédiction de la dimension, dûe à McEneaney, a été donnée dans [28], [50].

English version

The PhD work of Zheng Qu, supervised by S. Gaubert and S. Tang, dealt with the development of tropical methods in approximate dynamic programming [12].

The max-plus methods lead to approach the value function of an optimal control or zero-sum game problem by a supremum of a finite number of quadratic forms, see in particular [113]. We are interested here in the theoretical analysis (complexity) of this class of methods, as well as of their improvement. In certain cases, the quadratic forms are propagated by the flows of generalized Riccati equations. In order to perform an error analysis, we need to use some contraction properties of the Riccati flow, for certain known metrics on the space of positive matrices, like Thompson's metric. The latter is nothing but $d_T(A, B) = \| \log \operatorname{spec}(A^{-1}B) \|_\infty$, where spec denotes the sequence of eigenvalues of a matrix, and $\log$ is understood entrywise.

This led us to study the general problem of computing the contraction rate of an order-preserving flow on a cone, with respect to Thompson's metric. Indeed, the contraction properties of the standard Riccati flow are known (theorem of Bougerol for the invariant Riemanian metric, of Wojtowski for the Thompson's metric), but the proof of these properties (based on symplectic semigroups) does not carry over to generalized Riccati equations.

We gave in [114],[28] a general explicit formula for the contraction rate with respect to Thompson's metric of an order-preserving flow, involving only the generator of the flow and its derivative. We applied in particular this result to a generalized Riccati equation, associated to stochastic optimal control problems with a quadratic cost and a bilinear dynamics (presence of a bilinear term between the control and the noise). We showed that in this case, the Thompson's metric is the only invariant Finsler metric in which the generalized Riccati flow is nonexpansive, and we characterized the local contraction rate of this flow.

Z. Qu has applied these results in [28], [50] to the analysis of a method of reduction of the curse of dimensionality, introduced by McEneaney.

6.4.5. *Points fixes d'applications monotones homogènes et jeux à somme nulle/Fixed points of order preserving homogeneous maps and zero-sum games*
Participants: Marianne Akian, Stéphane Gaubert, Antoine Hochart.

Les opérateurs de Shapley sont les opérateur de programmation dynamique pour des jeux à somme nulle, ce sont précisément les opérateurs qui préservent l'ordre et commutent avec l'addition d'une constante. Le travail de M2 d'Antoine Hochart a traité d'une sous-classe d'opérateurs de Shapley, qui commutent en outre avec la multiplication par une constante positive. Nous les appellerons ici *sans-paiement*, car ils apparaissent dans des classes de jeux où les paiements instantanés sont nuls - le paiement a lieu seulement le dernier jour (*recursive games*). Ils apparaissent aussi dans l'étude structurelle de familles paramétriques de jeux répétés avec espace d'état fini et information parfaite, si l'on suppose par exemple que les probabilités de transitions sont fixées, mais que les paiements sont des paramètres. À toute famille paramétrique de jeux est associée un opérateur sans paiements et les points fixes de ce dernier sont précisément les vecteurs de paiement moyen réalisables. Un problème de base consiste à vérifier si un opérateur sans paiement n'a que des points fixes triviaux (réduits à des multiples du vecteur unité), et si possible, de déterminer des caractéristiques plus précises de l'ensemble des points-fixes, par exemple, savoir s'il existe un point fixe d'argmin donné. Le premier problème est connu être co-NP-complet, même pour des jeux déterministes. Nous montrons cependant que le second problème (point fixe d'argming prescrit) peut être résolu en temps polynomial. La preuve repose sur la construction d'une correspondance de Galois entre les faces d'un hypercube qui sont invariantes par l'opérateur, ainsi que sur une réduction à un problème d'accessibilité dans un hypergraphe orienté.

English version

Shapley operators are the dynamic programming operators of zero-sum stochastic games, they can be characterized as order preserving maps commuting with the addition of a constant. The M2 work of Antoine Hochart has dealt with a subclass of Shapley operators which are characterized by the property of commuting with the multiplication by a positive constant. We call them *payment-free*, as they arise in the study of *recursive games*, in which the payment only occurs when the game stops. They also arise in the study of structural properties of parametric mean payoff games (the transition probabilities are fixed, not the transition payoffs) with finite action spaces and perfect information: their fixed point set can be shown to give all the possible mean payoff vectors of such games. A basic problem is to check whether the fixed point set of such an operator is trivial (reduced to the multiples of the unit vector), and more precisely to determine its characteristics, for instance decide whether there is a fixed point with a prescribed argmin. The former problem is already known to be co-NP-complete, even for deterministic games. We showed however that the latter can be solved in polynomial time. The proof relies on the construction of a Galois connection between faces of the hypercube that are invariant by the operator, and on a reduction to a reachability problem in a directed hypergraph.

6.5. Applications

6.5.1. Introduction

Nous présentons maintenant plusieurs travaux de nature appliquée, touchant à des domaines variés, dans lesquels nous exploitons certaines des techniques mathématiques présentées précédemment, et particulièrement celles qui relèvent de la théorie de Perron-Frobenius non-linéaire et de la convexité tropicale. Ces applications utilisent aussi des techniques d'algèbre linéaire ou d'optimisation convexe.

English version

In this section, we describe several applied works in which we use some of the theoretical tools developed by the team, including non-linear Perron-Frobenius theory and tropical convexity. Some of these applications also make an intensive use of linear algebraic and convex programming methods.

6.5.2. Propriétés des valeurs propres de Perron et de Floquet, et application en chronothérapeutique/Properties of Perron and Floquet eigenvalue, with an application to chronotherapeutics

Participants: Frédérique Billy [Projet BANG, Inria], Jean Clairambault [Projet BANG, Inria], Olivier Fercoq, Stéphane Gaubert, Thomas Lepoutre [Projet BANG puis DRACULA, Inria].

On s'intéresse à des modèles de systèmes dynamiques monotones structurés en âge représentant la croissance de populations de cellules (saines ou tumorales), à la suite de travaux de Clairambault et Perthame. Il s'agit de comprendre l'influence du contrôle circardien sur la croissance des cellules. Dans le cas stationnaire, le taux de croissance est représenté par une valeur propre de Perron. Dans le cas périodique, il s'agit d'une valeur propre de Floquet. Les travaux [39], [73], [72] portent sur l'identification de ces modèles ainsi que sur un problème de contrôle thérapeutique, consistant à minimiser le taux de croissance des cellules tumorales sous une contrainte de non-toxicité du traitement (maintien d'une population de cellules saines). Ce travail s'appuie en particulier sur un algorithme d'optimisation de la valeur propre de Perron d'une matrice développé par Fercoq dans un autre contexte [104].

Un développement récent de ce travail peut être trouvé dans [39]. Un travail théorique sur ce type de modèles est présenté dans [46].

English version

We study monotone dynamical systems representing the growth of cells (healthy or tumoral), following a work of Clairambault and Perthame. The goal is to understand how the circadian control influences the growth of cells. In the case of stationnary monotone systems, this growth is measured by the Perron root. In the time periodic case, this Perron root is replaced by a Floquet multiplier.

The works [39], [73], [72] deal with the identification of these models, together with a therapeutic control problem, consisting in minimizing the growth rate of tumoral cells, under a non-toxicity constraint (preserving the population of healthy cells). This works relies in particular on a fast algorithm to optimize the Perron eigenvalue of a matrix, developed by Fercoq in a different context [104].

A recent development of this work can be found in [39]. A theoretical work on this kind of models has been presented in [46].

6.5.3. Preuve formelle d'inégalités non-linéaires/Formal proofs of non-linear inequalities

Participants: Xavier Allamigeon, Stéphane Gaubert, Victor Magron, Benjamin Werner [LIX].

La thèse de Victor Magron [11], dirigée par Benjamin Werner, codirigée par Stéphane Gaubert et Xavier Allamigeon, a porté sur la certification de bornes inférieures de fonctions multivariées à valeurs réelles, définies par des expressions semi-algébriques ou transcendantes, et sur la preuve de validité de celles-ci au moyen de certificats dans l'assistant de preuves Coq.

De nombreuses inégalités de cette nature apparaissent notamment dans la preuve par Thomas Hales de la conjecture de Kepler. Voici un exemple typique d'inégalité à prouver.

LEMME 9922699028 FLYSPECK. *Soit K, $\Delta\mathbf{x}$, l, t et f définis comme suit:*

$$
\begin{aligned}
K &:= [4, 6.3504]^3 \times [6.3504, 8] \times [4, 6.3504]^2 \ , \\
\Delta\mathbf{x} &:= x_1 x_4(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6) \\
&\quad + x_2 x_5(x_1 - x_2 + x_3 + x_4 - x_5 + x_6) \\
&\quad + x_3 x_6(x_1 + x_2 - x_3 + x_4 + x_5 - x_6) \\
&\quad - x_2 x_3 x_4 - x_1 x_3 x_5 - x_1 x_2 x_6 - x_4 x_5 x_6 \ , \\
l(\mathbf{x}) &:= -\pi/2 + 1.6294 - 0.2213(\sqrt{x_2} + \sqrt{x_3} + \sqrt{x_5} + \sqrt{x_6} - 8.0) \\
&\quad + 0.913(\sqrt{x_4} - 2.52) + 0.728(\sqrt{x_1} - 2.0) \ , \\
t(\mathbf{x}) &:= \arctan \frac{\partial_4 \Delta\mathbf{x}}{\sqrt{4 x_1 \Delta\mathbf{x}}} \ , \\
f(\mathbf{x}) &:= l(\mathbf{x}) + t(\mathbf{x}) \ .
\end{aligned}
$$

Alors, $\forall \mathbf{x} \in K, f(\mathbf{x}) \geq 0$.

On s'est donc intéressé à des fonctions non-linéaires, faisant intervenir des opérations semi-algébriques ainsi que des fonctions transcendantes univariées (cos, arctan, exp, etc).

De manière classique, on peut approcher les fonctions transcendantes qui interviennent de la sorte par des polynômes, ce qui permet de se ramener à des problèmes d'optimisation semi-algébriques, que l'on peut résoudre par des techniques de sommes de carrés creuses conduisant à des problèmes SDP. Cependant, en pratique, cette approche est limitée par la taille des SDP à résoudre, qui croît rapidement avec le degré des approximations polynomiales.

Dans ce travail de thèse, on a développé une méthode alternative, qui consiste a borner certains des constituants de la fonction non-linéaire par des suprema de formes quadratiques dont les Hessiens sont judicieusement choisis. On reprend donc ici l'idée des approximations "max-plus" initialement introduites en contrôle optimal, en s'appuyant sur des techniques d'interprétation abstraite (généralisation non-linéaire de la méthode des gabarits de Manna et al.). Ainsi, on obtient une nouvelle technique d'optimisation globale, basée sur les gabarits, qui exploite à la fois la precision des sommes de carrés et la capacité de passage à l'échelle des méthodes d'abstraction.

L'implémentation de ces méthodes d'approximation a abouti à un outil logiciel : NLCertify. Cet outil génère des certificats à partir d'approximations semi-algébriques et de sommes de carrés. Son interface avec Coq permet de bénéficier de l'arithmétique certifiée disponible dans l'assistant de preuves, et ainsi d'obtenir des estimateurs et des bornes valides pour chaque approximation.

Les performances de cet outil de certification ont été démontrées sur divers problèmes d'optimisation globale ainsi que sur des inégalités essentiellement serrées qui interviennent dans la preuve de Hales (projet Flyspeck).

Ce travail est exposé dans [25], [26].

English version

The PhD work of Victor Magron [11], supervised by Benjamin Werner, and cosupervised by Stéphane Gaubert and Xavier Allamigeon, dealt with the certification of lower bounds for multivariate functions, defined by semi-algebraic or transcendental expressions, and their correctness proof through certificates checked in the Coq proof assistant.

Many inequalities of this kind appear in particular in the proof by Thomas Hales of Kepler's conjecture. Here is a typical example of inequality.

LEMMA 9922699028 FLYSPECK. *Let K, $\Delta\mathbf{x}$, l, t and f be defined as follows:*

$$
\begin{aligned}
K &:= [4, 6.3504]^3 \times [6.3504, 8] \times [4, 6.3504]^2 \ , \\
\Delta\mathbf{x} &:= x_1 x_4(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6) \\
&\quad + x_2 x_5(x_1 - x_2 + x_3 + x_4 - x_5 + x_6) \\
&\quad + x_3 x_6(x_1 + x_2 - x_3 + x_4 + x_5 - x_6) \\
&\quad - x_2 x_3 x_4 - x_1 x_3 x_5 - x_1 x_2 x_6 - x_4 x_5 x_6 \ , \\
l(\mathbf{x}) &:= -\pi/2 + 1.6294 - 0.2213(\sqrt{x_2} + \sqrt{x_3} + \sqrt{x_5} + \sqrt{x_6} - 8.0) \\
&\quad + 0.913(\sqrt{x_4} - 2.52) + 0.728(\sqrt{x_1} - 2.0) \ , \\
t(\mathbf{x}) &:= \arctan \frac{\partial_4 \Delta\mathbf{x}}{\sqrt{4 x_1 \Delta\mathbf{x}}} \ , \\
f(\mathbf{x}) &:= l(\mathbf{x}) + t(\mathbf{x}) \ .
\end{aligned}
$$

Then, $\forall \mathbf{x} \in K, f(\mathbf{x}) \geq 0$.

Thus, we considered non-linear functions, defined in terms of semi-algebraic operations and univariate transcendental functions (cos, arctan, exp, etc).

Such transcendental functions can be classically approximated by polynomials, which leads to semi-algebraic optimization problems, which can be solved by sparse sum of squares techniques leading to SDP formulations. However, in practice, this approach is limited by the growth of the size of the SDP instances to be solved, whichs grows quickly with the degree of polynomial approximations.

In this PhD, we developed an alternative method, which consists in bounding some constituents of the non-linear function to be optimized by suprema of quadratic forms with well chosen Hessians. This is based on the idea of "maxplus approximation" initially introduced in optimal control, and also, on abstract interpretation (the template method introduced by Manna et al. in static analysis). In this way, we end up with a new global optimization technique, which takes advantage of the precision of sum of squares and of the scalability of abstraction methods.

These methods have been implemented in a software tool: NLCertify. This tool generates certificates from semi-algebraic and sum of square certificates. Its interface with Coq allows one to take benefit of the certified arithmetics available in this proof assistant, and so, to obtain estimators and valid bounds for each approximation.

The performances of this certification tool have been shown on several global optimization problems from the litterature, as well as on essentially tight inequalities taken from Hales' proof (Flyspeck project).

This work is presented in [25], [26].

6.5.4. *Vérification de systèmes temps-réels/Verification of real-time systems*
Participants: Xavier Allamigeon, Uli Fahrenberg [IRISA], Stéphane Gaubert, Ricardo Katz [Conicet], Axel Legay [IRISA].

Dans [141], Lu, Madsen, Milata, Ravn, Fahrenberg et Larsen ont montré que les polyèdres tropicaux peuvent être utilisés dans le cadre de l'analyse d'accessibilité d'automates temporisés. En effet, les polyèdres tropicaux expriment naturellement des invariants non-convexes, qui sont en fait des disjonctions d'invariants fournis par des DBM (*difference bound matrices*). A ce titre, les polyèdres tropicaux devraient permettre de réduire le nombre de disjonctions réalisées pendant l'analyse d'automates temporisés. Une limitation importante de cette approche est cependant que les polyèdres tropicaux sont topologiquement fermés, et qu'ils ne peuvent donc pas exprimer de contraintes d'inégalités strictes. Ces dernières sont néanmoins fondamentales dans l'analyse de systèmes temps-réels.

Nous avons donc développé dans [44] une généralisation des polyèdres tropicaux permettant d'exprimer des contraintes mixtes, *i.e.* strictes ou larges. Notre approche repose sur l'utilisation d'inégalités tropicales linéaires à coefficients dans un (quotient du) semi-anneau de germes affines. Afin de réaliser des opérations sur cette nouvelle classe de polyèdres tropicaux, nous avons défini deux nouveaux algorithmes. Le premier est un analogue tropical de l'élimination de Fourier-Motzkin. Celle-ci s'applique plus généralement à des systèmes d'inégalités linéaires sur des semi-anneaux idempotents et totalement ordonnés. Le second algorithme permet de tester si un système de contraintes mixtes admet une solution. Nous montrons en effet que ce problème est équivalent en temps polynomial à la résolution d'un problème de jeux déterministes à somme nulle. Ces deux contributions nous permettent de définir les primitives requises pour l'analyse d'accessibilité d'automates temporisés.

English version

Lu, Madsen, Milata, Ravn, Fahrenberg and Larsen have shown in [141] that tropical polyhedra can be applied to the reachability analysis of timed automata. Indeed, tropical polyhedra naturally express non-convex invariants, which correspond to disjunctions of invariants provided by DBM (*difference bound matrices*). Consequently, tropical polyhedra should allow to reduce the number of disjunctions arising during the analysis of timed automata. An important limitation of this approach is that tropical polyhedra are topologically closed, and thus they cannot express strict inequality constraints. However, such constraints plays an important role in the analysis of real-time systems.

As a result, we have developed in [44] a generalization of tropical polyhedra, in order to express mixed constraints, *i.e.* strict or loose ones. Our approach relies on tropical linear inequalities with coefficients in a (quotient of) the semiring of affine germs. In order to perform operations on this new class of polyhedra, we have introduced two new algorithms. The first one is a tropical analog of Fourier-Moztkin elimination. In fact, it applies more generally to systems of linear inequalities over totally ordered and idempotent semirings. The second algorithm allows to test the feasability of a mixed constraint system. We indeed show that this problem is polynomial-time equivalent to solving mean payoff games. These two contributions allow to define the primitives required by the reachability analysis of timed automata.

<p style="text-align:center">**MEXICO Project-Team**</p>

6. New Results

6.1. Diagnosis

- For non-diagnosable discrete event systems, *active* diagnosis aims at synthesizing a partial-observabion based control for the system in order to make it diagnosable. While some solutions had already been proposed for the active diagnosis problem, their complexity remained to be improved. In [40], we solved both the active diagnosability decision problem and the active diagnoser synthesis problem, proving that (1) our procedures are optimal w.r.t. to computational complexity, and (2) the memory required for the active diagnoser produced by the synthesis is minimal. Furthermore, focusing on the minimal delay before detection, we establish that the memory required for any active diagnoser achieving this delay may be highly greater than the previous one. So we refine our construction to build with the same complexity and memory requirement an active diagnoser that realizes a delay bounded by twice the minimal delay. An extension to *probabilistic* systems has been accepted to *FoSSaCS 2014.*

- In [41], we present a methodology for fault diagnosis in concurrent, partially observable systems with additional fairness constraints. In this *weak* diagnosis, one asks whether a concurrent chronicle of observed events allows to determine that a non-observable fault will inevitably occur, sooner or later, on any maximal system run compatible with the observation. The approach builds on strengths and techniques of unfoldings of safe Petri nets, striving to compute a compact prefix of the unfolding that carries sufficient information for the diagnosis algorithm. Our work extends and generalizes the unfolding-based diagnosis approaches by Benveniste et al. as well as Esparza and Kern. Both of these focused mostly on the use of sequential observations, in particular did not exploit the capacity of unfoldings to reveal inevitable occurrences of concurrent or future events studied by Balaguer et al. [19]. Our diagnosis method captures such indirect, revealed dependencies. We develop theoretical foundations and an algorithmic solution to the diagnosis problem, and present a SAT solving method for practical diagnosis with our approach. The algorithms to check diagnosability of concurrent systems are usually performed by local diagnoses of twin plant communicating with each other, directly or through a co- ordinator, and by that means pooling together the observations. Parallel analysis of diagnosability [43] takes advantage of the distribution of the system allowing to decide the diagnosability of the whole system in terms of the diagnosability of smaller systems.

6.2. Testing for Concurrent Systems

6.2.1. *Model Based Testing with Labeled Event Structures*

In [52], we have developed a complete testing framework for concurrent systems, which included the notions of test suites and test cases. We studied what kind of systems are testable in such a framework, and we have proposed sufficient conditions for obtaining a complete test suite as well as an algorithm to construct a test suite with such properties. However complete test suites are usually infinite in practice. In [44] (and a submitted journal version), we have proposed several testing criteria based on dedicated notions of complete prefixes that selects a manageable test suite together with a coverable criterion that allows to compare them.

6.3. Petri Nets

6.3.1. *A Modular Approach for Reusing Formalisms in Verification Tools of Concurrent Systems*

Over the past two decades, numerous verification tools have been successfully used for verifying complex concurrent systems, modelled using various formalisms. However, it is still hard to coordinate these tools since they rely on such a large number of formalisms. Having a proper syntactical mechanism to interrelate them through variability would increase the capability of effective integrated formal methods. In [28], we propose a modular approach for defining new formalisms by reusing existing ones and adding new features and/or constraints. Our approach relies on standard XML technologies; their use provides the capability of rapidly and automatically obtaining tools for representing and validating models. It thus enables fast iterations in developing and testing complex formalisms. As a case study, we applied our modular definition approach on families of Petri nets and timed automata.

6.3.2. *Computation of summaries using net unfoldings*

In [38], we study the following summarization problem: given a parallel composition A = A1 ∥ ... ∥ An of labelled transition systems communicating with the environment through a distinguished component Ai, efficiently compute a summary Si such that E ∥ A and E ∥ Si are trace-equivalent for every environment E. While Si can be computed using elementary automata theory, the resulting algorithm suffers from the state-explosion problem. We present a new, simple but subtle algorithm based on net unfoldings, a partial-order semantics, give experimental results. Our algorithm can also handle divergences and compute weighted summaries with minor modifications.

6.3.3. *Complexity Analysis of Continuous Petri Nets*

At the end of the eighties, continuous Petri nets were introduced for: (1) alleviating the combinatory explosion triggered by discrete Petri nets and, (2) modelling the behaviour of physical systems whose state is composed of continuous variables. Since then several works have established that the computational complexity of deciding some standard behavioural properties of Petri nets is reduced in this framework. In [39], we first establish the decidability of additional properties like boundedness and reachability set inclusion. We also design new decision procedures for the reachability and lim-reachability problems with a better computational complexity. Finally we provide lower bounds characterising the exact complexity class of the boundedness, the reachability, the deadlock freeness and the liveness problems.

6.3.4. *Contextual Merged Processes*

In [45], we integrate two compact data structures for representing state spaces of Petri nets: merged processes and contextual prefixes. The resulting data structure, called contextual merged processes (CMP), combines the advantages of the original ones and copes with several important sources of state space explosion: concurrency, sequences of choices, and concurrent read accesses to shared resources. In particular, we demonstrate on a number of benchmarks that CMPs are more compact than either of the original data structures. Moreover, we sketch a polynomial (in the CMP size) encoding into SAT of the model-checking problem for reachability properties.

6.3.5. *A Canonical Contraction for Safe Petri Nets*

Under maximal semantics, the occurrence of an event a in a concurrent run of an occurrence net may imply the occurrence of other events, not causally related to a, in the same run. In recent works, we have formalized this phenomenon as the *reveals* relation, and used it to obtain a contraction of sets of events called *facets* in the context of occurrence nets. In [36], we extend this idea to propose a canonical contraction of general safe Petri nets into pieces of partial-order behaviour which can be seen as "macro-transitions" since all their events must occur together in maximal semantics. On occurrence nets, our construction coincides with the facets abstraction. Our contraction preserves the maximal semantics in the sense that the maximal processes of the contracted net are in bijection with those of the original net.

6.4. Composition

6.4.1. *Specification of Asynchronous Component Systems with Modal I/O-Petri Nets*

In collaboration with Professor Rolf Hennicker from LMU and M.H. Møller, a PhD student from Aalborg University, we have studied the asynchronous composition of systems where the internal channels remain observable. In [42], we have modelled such systems by Petri nets enlarged with communication channels, we have defined several channel properties and shown these properties are compositional, and proved their decidability. In TGC 2013 (not yet in HAL), we have extended the previous models with modalities *must* and *may* "à la Larsen" and generalized most of the results in this framework.

6.4.2. Bounding models families for performance evaluation in composite Web services

One challenge of composite Web service architectures is the guarantee of the Quality of Service (QoS). Performance evaluation of these architectures is essential but complex due to synchronizations inside the orchestration of services. In (ADD WHEN IN HAL), we propose methods to automatically derive from the original model a family of bounding models for the composite Web response time. These models allow to find the appropriate trade-off between accuracy of the bounds and the computational complexity. The numerical results show the interest of our approach w.r.t. complexity and accuracy of the response time bounds.

6.5. Stochastic Systems

6.5.1. Simulation-based Verification of HASL (Hybrid Automata Stochastic Logic) Formulas for Stochastic Symmetric Nets

The Hybrid Automata Stochastic Logic (HASL) has been recently defined as a flexible way to express classical performance measures as well as more complex, path-based ones (generically called "HASL formulas"). The considered paths are executions of Generalized Stochastic Petri Nets (GSPN), which are an extension of the basic Petri net formalism to define discrete event stochastic processes. The computation of the HASL formulas for a GSPN model is demanded to the COSMOS tool, that applies simulation techniques to the formula computation. Stochastic Symmetric Nets (SSN) are an high level Petri net formalism, of the colored type, in which tokens can have an identity, and it is well known that colored Petri nets allow one to describe systems in a more compact and parametric form than basic (uncolored) Petri nets. In [27], we propose to extend HASL and COSMOS to support colors, so that performance formulas for SSN can be easily defined and evaluated. This requires a new definition of the logic, to ensure that colors are taken into account in a correct and useful manner, and a significant extension of the COSMOS tool.

6.5.2. Steady-state control problem for Markov decision processes

We address in (ADD CITATION WHEN IN HAL) a control problem for probabilistic models in the setting of Markov decision processes (MDP). We are interested in the steady-state control problem which asks, given an ergodic MDP M and a distribution δ, whether there exists a (history-dependent randomized) policy π ensuring that the steady-state distribution of M under $\imath$ is exactly δ. We first show that stationary randomized policies suffice to achieve a given steady-state distribution. Then we infer that the steady-state control problem is decidable for MDP, and can be represented as a linear program which is solvable in PTIME. This decidability result extends to labeled MDP (LMDP) where the objective is a steady-state distribution on labels carried by the states, and we provide a PSPACE algorithm. We also show that a related steady-state language inclusion problem is decidable in EXPTIME for LMDP. Finally, we prove that if we consider MDP under partial observation (POMDP), the steady-state control problem becomes undecidable.

6.6. Timed Systems

6.6.1. Back in Time Petri Nets

The time progress assumption is at the core of the semantics of real-time formalisms. It is also the major obstacle to the development of partial-order techniques for real-time distributed systems since the events are ordered both by causality and by their occurrence in time. Anyway, extended free choice safe time Petri nets (TPNs) were already identified as a class where partial order semantics behaves well. In [37], we show that, for this class, the time progress assumption can even be dropped (time may go back in case of concurrency), which establishes a nice relation between partial-order semantics and time progress assumption.

6.6.2. Expressiveness of Timed Models

In coopération with Nantes and UPMC, an in-depth study of the expressiveness of time Petri nets was completed [20]. With roughly the same partners, we have extended th ITA (Interrupt Timed Automata) by parametrizing both guards and clock rates while preserving the decidability results (RP 2013, not yet in HAL).

6.7. Weighted Systems

6.7.1. Specification and Verification of Quantitative Properties via Expressions, Logics, and Automata

Alongside boolean properties, automatic verification of *quantitative* properties such as lifespan of an equipment, energy consumption of an application or reliability of a program is gaining importance rapidly. In the thesis [14] and the articles [32], [14] , several weight-enabled formalisms for specification of such properties were examined, including denotational ones such as regular expressions, first-order logic with transitive closure, or temporal logics, as well as more operational ones such as navigating automata, possibly extended with pebbles. A unified framework of graph structures allows to compare these formalisms with respect to expressiveness, using efficient translations from denotational to operational formalisms. Several decidability and complexity results for the algorithmic questions that arise were obtained, depending on the underlying semiring from which weights are chosen, and on the structures (words, trees, ...) considered.

6.8. Dynamic Communicating Systems

6.8.1. Specification and Verification of Dynamic Message-Passing Systems

In [31], we study dynamic communicating automata (DCA), an extension of classical communicating finite-state machines that allows for dynamic creation of processes. The behavior of a DCA can be described as a set of message sequence charts (MSCs). While DCA serve as a model of an implementation, we propose branching high-level MSCs (bHMSCs) on the specification side. Our focus is on the implementability problem: given a bHMSC, can one construct an equivalent DCA? As this problem is undecidable, we introduce the notion of executability, a decidable necessary criterion for implementability. We show that executability of bHMSCs is EXPTIME-complete. We then identify a class of bHMSCs for which executability effectively implies implementability.

6.9. Concurrent Recursive Programs

6.9.1. The Complexity of Model Checking Concurrent Recursive Programs

In [34], we consider the linear-time model checking problem for boolean concurrent programs with recursive procedure calls. While sequential recursive programs are usually modeled as pushdown automata, concurrent recursive programs involve several processes and can be naturally abstracted as pushdown automata with multiple stacks. Their behavior can be understood as words with multiple nesting relations, each relation connecting a procedure call with its corresponding return. To reason about multiply nested words, we consider the class of all temporal logics as defined in the book by Gabbay, Hodkinson, and Reynolds (1994). The unifying feature of these temporal logics is that their modalities are defined in monadic second-order (MSO) logic. In particular, this captures numerous temporal logics over concurrent and/or recursive programs that have been defined so far. Since the general model checking problem is undecidable, we restrict attention to phase bounded executions as proposed by La Torre, Madhusudan, and Parlato (LICS 2007). While the MSO model checking problem in this case is non-elementary, our main result states that the model checking (and satisfiability) problem for all MSO-definable temporal logics is decidable in elementary time. More precisely, it is solvable in $(n + 2)$-EXPTIME where n is the maximal level of the MSO modalities in the monadic quantifier alternation hierarchy. We complement this result and provide, for each level n, a temporal logic whose model checking problem is n-EXPSPACE-hard.

6.9.2. *Model Checking Concurrent Recursive and Communicating Programs via Split-Width*

The work described in the following was done by Aiswarya Cyriac in collaboration with Paul Gastin and K. Narayan Kumar, and it is part of Aiswarya Cyriac's PhD thesis, which has recently been defended. It is a generalisation of our CONCUR'12 paper where split-width is introduced to address the decidability of MSO specifications for multi-pushdown systems.

We consider generic systems which incorporate shared-variable communication and communication via channels. We are considering physically distributed machines which communicate via (possibly several) reliable first-in-first-out queues. Each of these machines are capable of running potentially recursive multi-threaded programs. These programs within a machine use shared variable for communication. Such a machine consisting of a set of threads communicating by shared memory can be formally modelled as a multi-pushdown system. Thus we have a network of multi-pushdown systems communicating via FIFO queues. Moreover, these programs may use stacks and queues as data-structures to aid their local computation. We call such a system a system of concurrent processes with data-structures (CPDS).

We introduce and study a new technique called split-width for the under-approximate verification of CPDS. This parameter is based on simple shuffle and merge operations and gives us a divide-conquer-way to prove the bound of languages. When parametrised by a bound on split-width, we obtain decidability for various verification problems. We provide a uniform decision procedure for various verification problems with optimal complexities.

We expose the power of split-width in several ways. We show that our simple algebra is powerful enough to capture any class of CPDS which admits decidability for MSO model checking, and yardstick graph metrics such as tree-width and clique-width. We also show that various restrictions well-studied in the literature for obtaining decidability of reachability for the particular cases of multi-pushdown systems and message passing systems admit a bound on split-width. In fact, we propose generic controllers which subsume many of these cases.

Distributed controller design amounts to designing a controller (which is another CPDS) which, when run sychronously with a system ensures bounded split-width. These controllers are distributed in nature and are independent of the system it is controlling. Thus such a controller respects the privacy of the system (by not reading their states, for instance). Moreover, thanks to split-width such a controlled system offers efficient (in most cases optimal) decision procedures for the verification of the controlled system. We propose a generic approach to define controllable classes of CPDS in terms of quotient graphs, which admit a "suitable" acyclicity restriction. We also give a generic controller for several of the classes definable in this framework. The controllers we propose are sound and complete for the respective class, meaning that they allow all and only the behaviours of this class. Moreover, our technique for proving the bound on split-width of the controlled systems is also generic and systematic, hence may easily extend to generalisations and other classes as well.

The decidability results for the controllable classes proposed in the thesis are new while they capture, as special cases, several restrictions studied in the literature like bounded phase, bounded scope, poly-forest topology etc.

<p style="text-align:center">**OAK Project-Team**</p>

6. New Results

6.1. Scalable and Expressive Techniques for the Semantic Web

The team has continued developing expressive models and scalable algorithms for exploiting Semantic Web data, in particular RDF graphs, as well as rich corpora consisting of Web documents with semantic annotations.

We have studied efficient algorithms for answering RDF queries in the presence of schema (or semantic) constraints such as described through the RDF Schema language. The difficulty here consists of efficiently taking into account the data that is implicitly present in the RDF database due to semantic constraints, and which needs to be reflected in query results. We have identified the expressive database fragment of RDF, which extends previously identified fragments of the RDF specification by allowing more expressive schema and queries, and provided novel efficient algorithms for answering Basic Graph Pattern queries (a popular dialect of the standard SPARQL query language) over RDF graphs pertaining to the RDF Database Fragment. Our query answering algorithms take advantage of the processing power of a relational database management system while also reflecting RDF semantics [25].

The ability to exploit large corpora of heterogeneous RDF data requires tools for analyzing RDF content through the lenses of a specific user perspective, or user need. Such tools are commonplace in the context of relational data management, where data warehousing is a well-developed area, but lack completely in the realm of RDF. We have proposed a novel framework for building and exploiting all-RDF data warehouses [33] and have implemented this framework in a proof-of-concept platform [32]. A main contribution of this work is to preserve RDF graph structure, heterogeneity, and rich semantics from the base data to the analytical schema and analytical schema instance. Thus, our proposal is the first to allow the analysis of rich Semantic Web (RDF) data while preserving its rich content and semantics. For more information on this project, see https://team.inria.fr/oak/warg/.

We have investigated the usage of semantics as a way to enrich, interconnect, and interpret rich corpora of Web data. In particular, within the XR project, we had proposed in prior work the XR (XML+RDF) data model which integrates XML documents and RDF triples treating both as first-class citizens. One particular use of XR is to annotate nodes in XML documents, by RDF triples which may for instance describe their properties or state how nodes are semantically related to some concept or to each other. In [18] we describe the data model and core query language, make a comprehensive analysis of query evaluation algorithms, and describe extensive experiments carried within a fully implemented platform, as part of the PhD thesis of J. Leblay [12]. The XR platform was put to task in an application context related to digital journalism, where an XR content warehouse is continuously enriched through document analysis and annotation. This scenario has lead to a software demonstration [24], [35] and a keynote tutorial [38]. In collaboration with A. Deutsch, we have extended the XR query language and provided query-view composition algorithms in [41].

6.2. Massively Distributed Data Management Systems

Our work on the AMADA platform has shown how the different sub-systems of a popular cloud platform (namely, Amazon Web Services, or AWS in short) can be harnessed to build scalable stores and query evaluation engines for XML and RDF data. In [23], we propose and compare several storage and indexing strategies within AWS, and show that they help reduce not only query evaluation time but also the monetary costs associated to the exploitation of the AWS-based store, since the index helps direct queries only to the subsets of the data likely to have results for the query. Thus, the total effort (and the costs charged by AWS) in relation to the processing of a given query are reduced. A similar study focused mostly on RDF data management appears as a book chapter [40]. More information can be found at http://cloak.saclay.inria.fr/research/amada/.

Semantic Web data collections, that is, RDF graphs, may be very voluminous since RDF natively enables connections between different RDF databases (which may have been produced independently and in ignorance of each other) through the usage of common URIs (resource identifiers) in two or more databases. To scale up to such large volumes, we have developed CliqueSquare, a novel platform for storing and querying RDF graphs in a MapReduce-based architecture such as Hadoop. We have described the storage and query algorithm in [34]. Our analysis of existing frameworks and algorithms for managing large RDF graphs in a highly distributed environment has lead to the tutorial [27].

Large-scale distributed processing of complex data was considered from a different perspective in our Delta project. Here, we considered the setting where one data source publishes new data items at a very high rate, and numerous clients subscribe to some of the updates by means of queries that must be matched by the published items. In this setting, the source may quickly become the bottleneck due to limitations in its capacity to match the published item against the subscription and/or to send the matching updates. We propose a fully automated approach for distributing the data dissemination effort across the network of subscribers, by identifying some which act as secondary data sources for others, in a peer-to-peer fashion. This distributed dissemination network is chosen so as to optimize a combination of overall dissemination costs and data propagation latency; since the space of options has daunting complexity, approximate algorithms involving Binary Integer Programming techniques were proposed in [20], [37], [42], and concluded in the PhD thesis of A. Katsifodimos [11].

6.3. Advanced Algorithms for Efficient XML processing

In 2013, several research works of the team focusing on advanced algorithms for processing XML data have been finalized and concluded through prestigious journal publications.

A first line of work concerned the usage of materialized views to speed up the evaluation of complex XML queries. In our previous work we had demonstrated that such views may bring up very significant speed-up factors of several orders of magnitude. However, materialized views need to be kept up to date when the underlying database changes. In [14] we have described efficient algorithms for updating materialized views expressed in a rich dialect of XQuery, the standard query language for XML.

A second class of work was concerned with XML static type analysis, in particular with the crucial problem of deciding XML type inclusion, that is: whether any XML tree of type τ_1 is also of type τ_2 where τ_1, τ_2 are XML types with interleaving and counting (currently adopted by main stream schema languages). For these types, inclusion is EXPSPACE-complete. We have defined and formally studied a quadratic subtype-checking algorithm for the case where the right-hand side type τ_2 meets some restrictions on symbol occurrences and the use of counting. These restrictions are often met by human-designed types, so our technique perfectly fits the needs of typical XML type-checking algorithms, which frequently require to check for inclusion a machine-generated subtype τ_1 against a human-defined supertype τ_2. Our approach has been validated by extensive experimental results [16]. In addition, we have devised and formally studied an alternative algorithm, still for the asymmetric case where τ_2 is restricted, based on structural, top-down analysis of types expression. This algorithm is almost linear: it has a linear-time backbone, and resorts to the above quadratic approach for some specific parts of the compared types. Our experiments show that this new algorithm is much faster than the quadratic one and that it typically runs in linear time, hence it can be used as a building block for a practical type-checking compiler for XML programs and queries [15].

Third, we have completed and concluded our work on type-based document projection for efficient XML data management. The idea here is to restrict XML documents, prior to evaluating a query over them, to only those parts of the document that the query actually needs to consult. We provide algorithms for determining such document parts and experimentally demonstrate the benefits of such techniques, in [13].

Finally, we have devised a system that is able to process both queries and updates on very large XML documents [22]. As observed in recent works, such very large documents are generated and processed in several contexts, in particular in those involving scientific data and logs. Our system supports a large fragment of XQuery and XUF (XQuery Update Facility). The system exploits dynamic and static partitioning to

distribute the processing load among the machines of a MapReduce cluster. The proposed technique applies when queries and updates are iterative, i.e., they iterate the same query/update operations on a sequence of subtrees of the input document. From our experience many real world queries and updates actually meet this property. Our partitioning technique is schema-less, as the presence of a user-supplied schema is not required; indeed, this technique only relies on path information extracted from the input query/update. Experiments conducted on a 8-machine Hadoop cluster have demonstrated that the system is able run both iterative queries and updates on quite large documents.

6.4. Data Transformation Management

With the increasing complexity of data processing queries, for instance in applications such as relational data analysis or integration of Web data (e.g., XML or RDF) comes the need to better manage complex data transformations. This includes systematically verifying, maintaining, and testing the transformations an application relies on. In this context, Oak has focused on verifying the semantic correctness of a declarative program that specifies a data transformation query, e.g., an SQL query. To this end, we have investigated how to leverage data provenance (the information of the origin of data and the query operators) for query debugging. More specifically, we developed and implemented novel algorithms to explain why data is missing from the result of a relational query. As opposed to our previous work, which produced explanations based on the available source data, our new algorithms return explanations based on query operators [31] or both [26].

6.5. Social Data Management

We considered top-k query answering in social tagging systems, also known as folksonomies, a problem that requires a significant departure from existing, socially agnostic techniques. In a network-aware context, one can and should exploit the social links, which can indicate how users relate to the seeker and how much weight their tagging actions should have in the result build-up. Our solutions addressed the main drawbacks of previous approaches. With respect to applicability and scalability, we avoid expensive and hardly updatable pre-computations of proximity values. With respect to efficiency, we show that our algorithm is instance optimal in the existing techniques. Our main results in this direction have been presented recently in [29], [28], [21].

PARIETAL Project-Team

6. New Results

6.1. Deformable Template estimation for joint anatomical and functional brain images

Participants: Bertrand Thirion [Correspondant], Hao Xu, Stéphanie Allassonnière.

Traditional analyses of Functional Magnetic Resonance Imaging (fMRI) use little anatomical information. The registration of the images to a template is based on the individual anatomy and ignores functional information; subsequently detected activations are not confined to gray matter (GM). In this work, we propose a statistical model to estimate a probabilistic atlas from functional and T1 MRIs that summarizes both anatomical and functional information and the geometric variability of the population. Registration and Segmentation are performed jointly along the atlas estimation and the functional activity is constrained to the GM, increasing the accuracy of the atlas.

More details can be found in [69].

6.2. Randomized parcellation-based inference

Participants: Gaël Varoquaux, Bertrand Thirion, Benoit Da Mota, Virgile Fritsch.

Neuroimaging group analyses are used to relate inter-subject signal differences observed in brain imaging with behavioral or genetic variables and to assess risks factors of brain diseases. The lack of stability and of sensitivity of current voxel-based analysis schemes may however lead to non-reproducible results. We introduce a new approach to overcome the limitations of standard methods, in which active voxels are detected according to a consensus on several random parcellations of the brain images, while a permutation test controls the false positive risk (see Fig. 3). Both on synthetic and real data, this approach shows higher sensitivity, better accuracy and higher reproducibility than state-of-the-art methods. In a neuroimaging-genetic application, we find that it succeeds in detecting a significant association between a genetic variant next to the COMT gene and the BOLD signal in the left thalamus for a functional Magnetic Resonance Imaging contrast associated with incorrect responses of the subjects from a Stop Signal Task protocol.

More details can be found in [55].

6.3. Group-level impacts of within- and between-subject hemodynamic variability in fMRI

Participants: Gaël Varoquaux, Solveig Badillo, Philippe Ciuciu [Correspondant].

Inter-subject fMRI analyses have specific issues regarding the reliability of the results concerning both the detection of brain activation patterns and the estimation of the underlying dynamics. Among these issues lies the variability of the hemodynamic response function (HRF), that is usually accounted for using functional basis sets in the general linear model context. Here, we use the joint detection-estimation approach (JDE) [76], [78], which combines regional nonparametric HRF inference with spatially adaptive regularization of activation clusters to avoid global smoothing of fMRI images (see Fig. 4). We show that the JDE-based inference brings a significant improvement in statistical sensitivity for detecting evoked activity in parietal regions. In contrast, the canonical HRF associated with spatially adaptive regularization is more sensitive in other regions, such as motor cortex. This different regional behavior is shown to reflect a larger discrepancy of HRF with the canonical model. By varying parallel imaging acceleration factor, SNR-specific region-based hemodynamic parameters (activation delay and duration) were extracted from the JDE inference. Complementary analyses highlighted their significant departure from the canonical parameters and the strongest between-subject variability that occurs in the parietal region, irrespective of the SNR value. Finally, statistical evidence that the fluctuation of the HRF shape is responsible for the significant change in activation detection performance is demonstrated using paired t-tests between hemodynamic parameters inferred by GLM and JDE.

../../../../projets/parietal/IMG/method.png

Figure 3. Overview of the randomized parcellation based inference framework on an example with few parcels. The variability of the parcels definition is used to obtain voxel-level statistics.

More details can be found in [49].

../../../../projets/parietal/IMG/solveig.png

*Figure 4. General sketch summarizing the HRF computation at the subject and group-levels in activated regions r.
Left: Position of the activation peak in r (here left motor cortex) given in mm in the Talairach space. Center:
Individual weighted HRF time course extraction. Right: Computation of the group average normalized HRF time
course with corresponding error bars ($\pm\sigma$).*

6.4. Mapping cognitive ontologies to and from the brain

Participants: Gaël Varoquaux [Correspondant], Bertrand Thirion, Yannick Schwartz.

Imaging neuroscience links brain activation maps to behavior and cognition via correlational studies. Due to the nature of the individual experiments, based on eliciting neural response from a small number of stimuli, this link is incomplete, and unidirectional from the causal point of view. To come to conclusions on the function implied by the activation of brain regions, it is necessary to combine a wide exploration of the various brain functions and some inversion of the statistical inference. Here we introduce a methodology for accumulating knowledge towards a bidirectional link between observed brain activity and the corresponding function. We rely on a large corpus of imaging studies and a predictive engine. Technically, the challenges are to find commonality between the studies without denaturing the richness of the corpus. The key elements that we contribute are labeling the tasks performed with a cognitive ontology, and modeling the long tail of rare paradigms in the corpus. To our knowledge, our approach is the first demonstration of predicting the cognitive content of completely new brain images. To that end, we propose a method that predicts the experimental paradigms across different studies (see Fig. 5).

More details can be found in [63].

6.5. Implications of Inconsistencies between fMRI and dMRI on Multimodal Connectivity Estimation

Participants: Gaël Varoquaux [Correspondant], Bertrand Thirion, Bernard Ng.

There is a recent trend towards integrating resting state functional magnetic resonance imaging (RS-fMRI) and diffusion MRI (dMRI) for brain connectivity estimation, as motivated by how estimates from these modalities are presumably two views reflecting the same underlying brain circuitry. In this work, we show on a cohort of 60 subjects that conventional functional connectivity (FC) estimates based on Pearson's correlation and anatomical connectivity (AC) estimates based on fiber counts are actually not that highly correlated for typical RS-fMRI (7 min) and dMRI (32 gradient directions) data. The FC-AC correlation can be significantly increased by considering sparse partial correlation and modeling fiber endpoint uncertainty, but the resulting FC-AC correlation is still rather low in absolute terms. We further exemplify the inconsistencies between FC and AC estimates by integrating them as priors into activation detection and demonstrating significant differences in their detection sensitivity. Importantly, we illustrate that these inconsistencies can be useful in fMRI-dMRI integration for improving brain connectivity estimation.

More details can be found in [61]. See also [60].

6.6. Extracting brain regions from rest fMRI with Total-Variation constrained dictionary learning

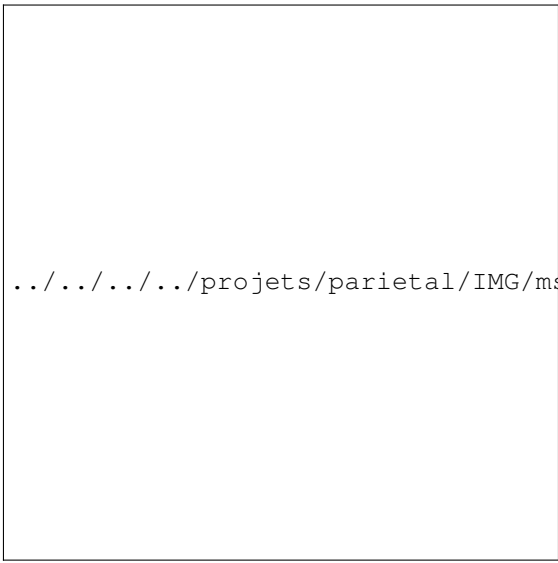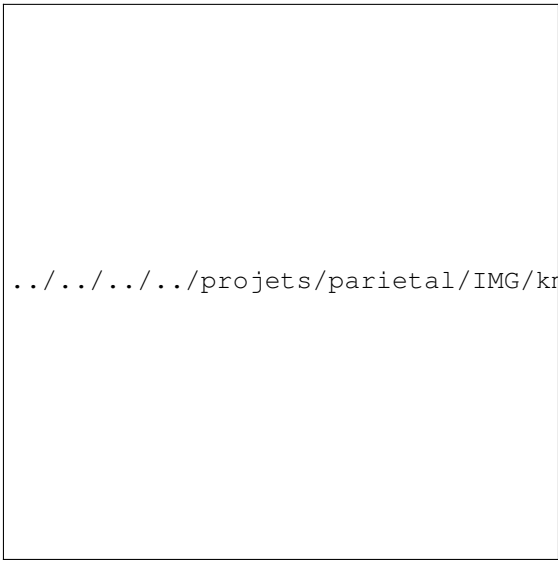**Participants:** Gaël Varoquaux [Correspondant], Alexandre Abraham.

Spontaneous brain activity reveals mechanisms of brain function and dysfunction. Its population-level statistical analysis based on functional images often relies on the de nition of brain regions that must summarize e ciently the covariance structure between the multiple brain networks. In this paper, we extend a network-discovery approach, namely dictionary learning, to readily extract brain regions. To do so, we intro duce a new tool drawing from clustering and linear decomposition methods by carefully crafting a penalty. Our approach automatically extracts regions from rest fMRI that better explain the data and are more stable across subjects than reference decomposition or clustering methods (see FIg. 6).

More details can be found in [47].

6.7. Cohort-level brain mapping: learning cognitive atoms to single out specialized regions

Participants: Gaël Varoquaux [Correspondant], Bertrand Thirion, Yannick Schwartz.

../../../../projets/parietal/IMG/forward_regions/stimulus_modality/IMG/reverse_regions

../../../../projets/parietal/IMG/forward_regions/response_modality/IMG/reverse_regions

../../../../projets/parietal/IMG/forward_regions/projets/parietal/IMG/reverse_regions

../../../../projets/parietal/IMG/kmeans_hard.png

../../../../projets/parietal/IMG/ward_hard.png

../../../../projets/parietal/IMG/msdl_hard.png

../../../../projets/parietal/IMG/msdl_unreg_hard

../../../../projets/parietal/IMG/smooth_msdl_hard.png

../../../../projets/parietal/IMG/icams_hard.png

Functional Magnetic Resonance Imaging (fMRI) studies map the human brain by testing the response of groups of individuals to carefully-crafted and contrasted tasks in order to delineate specialized brain regions and networks. The number of functional networks extracted is limited by the number of subject-level contrasts and does not grow with the cohort. Here, we introduce a new group-level brain mapping strategy to differentiate many regions reflecting the variety of brain network configurations observed in the population. Based on the principle of functional segregation, our approach singles out functionally-specialized brain regions by learning group-level functional profiles on which the response of brain regions can be represented sparsely. We use a dictionary-learning formulation that can be solved efficiently with on-line algorithms, scaling to arbitrary large datasets. Importantly, we model inter-subject correspondence as structure imposed in the estimated functional profiles, integrating a structure-inducing regularization with no additional computational cost. On a large multi-subject study, our approach extracts a large number of brain networks with meaningful functional profiles (see Fig. 7).

More details can be found in [66].

6.8. Identifying predictive regions from fMRI with TV-ℓ1 prior

Participants: Gaël Varoquaux [Correspondant], Bertrand Thirion, Alexandre Gramfort.

Decoding, i.e. predicting stimulus related quantities from functional brain images, is a powerful tool to demonstrate differences between brain activity across conditions. However, unlike standard brain mapping, it offers no guaranties on the localization of this information. Here, we consider decoding as a statistical estimation problem and show that injecting a spatial segmentation prior leads to unmatched performance in recovering predictive regions. Specifically, we use ℓ1 penalization to set voxels to zero and Total-Variation (TV) penalization to segment regions. Our contribution is two-fold. On the one hand, we show via extensive experiments that, amongst a large selection of decoding and brain-mapping strategies, TV+ℓ1 leads to best region recovery (see Fig. 8). On the other hand, we consider implementation issues related to this estimator. To tackle efficiently this joint prediction-segmentation problem we introduce a fast optimization algorithm based on a primal-dual approach. We also tackle automatic setting of hyper-parameters and fast computation of image operation on the irregular masks that arise in brain imaging.

More details can be found in [59].

6.9. Second order scattering descriptors predict fMRI activity due to visual textures

Participants: Michael Eickenberg, Bertrand Thirion [Correspondant], Alexandre Gramfort.

Second layer scattering descriptors are known to provide good classification performance on natural quasi-stationary processes such as visual textures due to their sensitivity to higher order moments and continuity with respect to small deformations. In a functional Magnetic Resonance Imaging (fMRI) experiment we present visual textures to subjects and evaluate the predictive power of these descriptors with respect to the predictive power of simple contour energy - the first scattering layer. We are able to conclude not only that invariant second layer scattering coefficients better encode voxel activity, but also that well predicted voxels need not necessarily lie in known retinotopic regions (see Fig. 9).

More details can be found in [56].

6.10. Bayesian Joint Detection-Estimation of cerebral vasoreactivity from ASL fMRI data

Participants: Thomas Vincent, Philippe Ciuciu [Correspondant].
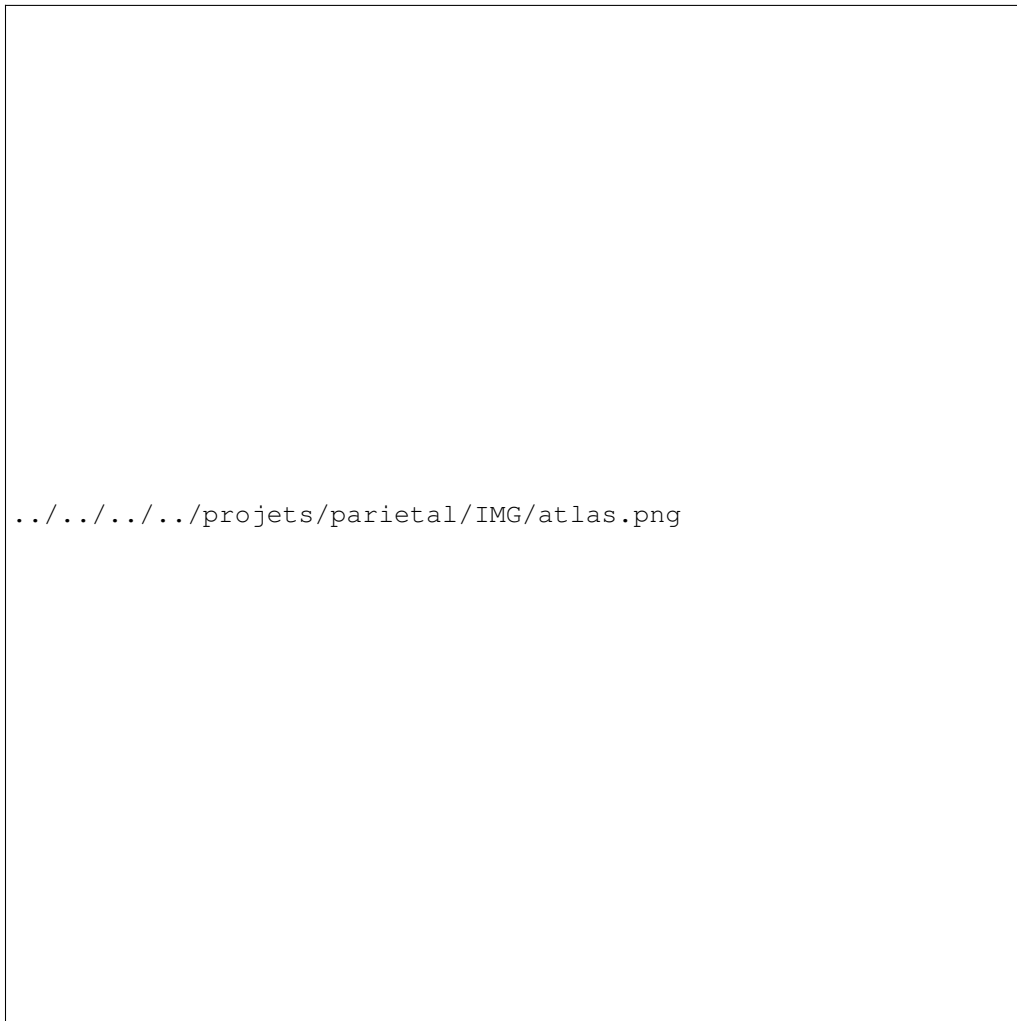
../../../../projets/parietal/IMG/atlas.png

Figure 7. (Left) A brain functional atlas can be conceptualized as a parcellation of the brain volume into overlapping networks, where each functional network is characterized by a profile of activation for a set of functional contrasts. (Right) Such an atlas can be learned by applying an adapted dictionary learning to a set of images that display the activation observed in different subjects for a (very large) set of cognitive tasks.
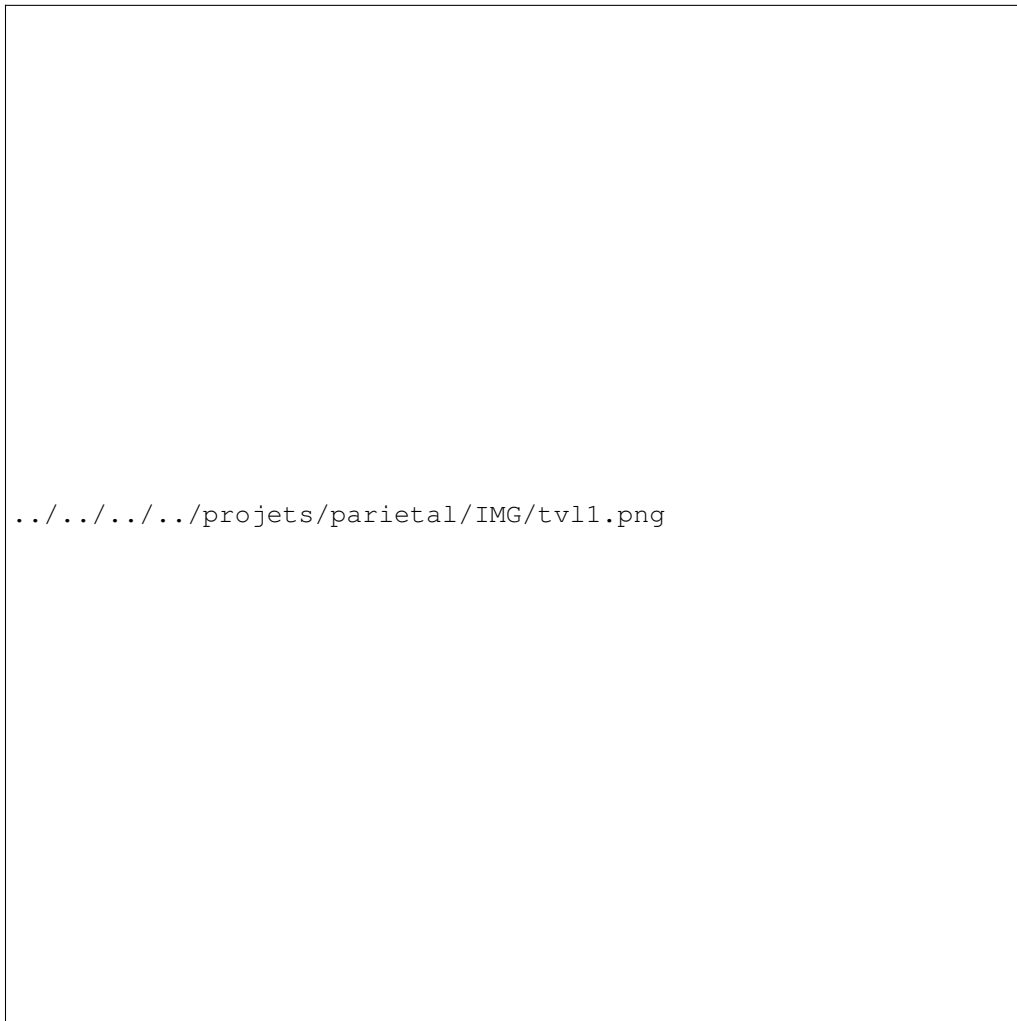
Figure 8. Results on fMRI data from (from left to right F-test, ElasticNet and TV-ℓ_1). The TV-ℓ_1 regularized model segments neuroscientificly meaningful predictive regions in agreement with univariate statistics while the ElasticNet yields sparse although very scattered non-zero weights.

../../../../projets/parietal/IMG/scattering.png

Figure 9. Some brain regions are better explained by using two scattering layers rather than one (middle). These regions are symetric across hemispheres, and are observed mostly in the dorsal stream of the visual cortex. An atlas of the visual areas (left and right) shows that the mai foci are found in the V1, V2, V3AB and IPS0 regions.

Although the study of cerebral vasoreactivity using fMRI is mainly conducted through the BOLD fMRI modality, owing to its relatively high signal-to-noise ratio (SNR), ASL fMRI provides a more interpretable measure of cerebral vasoreactivity than BOLD fMRI. Still, ASL suffers from a low SNR and is hampered by a large amount of physiological noise. The current contribution aims at improving the recovery of the vasoreactive component from the ASL signal. To this end, a Bayesian hierarchical model is proposed, enabling the recovery of perfusion levels as well as fitting their dynamics. On a single-subject ASL real data set involving perfusion changes induced by hypercapnia, the approach is compared with a classical GLM-based analysis. A better goodness-of-fit is achieved, especially in the transitions between baseline and hypercapnia periods. Also, perfusion levels are recovered with higher sensitivity and show a better contrast between gray- and white matter.

More details can be found in [68].

6. New Results

6.1. Substitution as Proof Compression

Participants: Lutz Straßburger, Novak Novakovic.

In previous work [58] we have shown how the calculus of structures can accommodate Tseitin extension without relying on the cut (or modus ponens). Thus, cut and extension can be studied independently as proof compression mechanisms. Another such proof compression mechanism is substitution. It has been shown by Cook, Reckhow, Krajíček and Pudlák that in the presence of cut, extension and substitution are equally powerful with respect to proof complexity. This year we succeeded in showing that this is also the case in the absence of cut. I.e., we have shown that the cut-free system with extension and the cut-free system with substitution p-simulate each other. This result is presented in [34].

6.2. Herbrand Confluence

Participants: Lutz Straßburger, Stefan Hetzl.

In the result on Herbrand confluence from last year [46], the endsequent of a proof had to be an existential sentence in prenex form. This year we were able to relax this restriction and to extend our result to arbitrary endsequent. This work has been published in [15].

6.3. Nested Sequents for Intuitionistic Modal Logics

Participant: Lutz Straßburger.

We present cut-free deductive systems without labels for the intuitionistic variants of the modal logics obtained by extending IK with a subset of the axioms d, t, b, 4, and 5. For this, we use the formalism of nested sequents, which allows us to give a uniform cut elimination argument for all 15 logic in the intuitionistic S5 cube. This work (published in [25]), is an improvement of the result on intuitionistic modal logic from 2011: the deductive systems the cut elimination proof are much simpler now.

6.4. First efforts at designing proof certificates

Participants: Hichem Chihani, Quentin Heath, Dale Miller, Fabien Renaud.

Work on the ERC Advance Grant ProofCert has progressed along two lines.

Given earlier work within the team [6], [7], there now exists a flexible and well understood concept of focused proof for classical and intuitionistic first-order logics. Chihani, Miller, and Renaud have been working to use that notion of proof as a means of providing flexible definition of *proof evidence* for those two logics. Initial results along those directions have been reported in the [19] and [20]. In those papers, several examples definitions of the semantics of *proof certificates* (formal documents providing the details of some proof evidence) are provided in such a way that a single, simple proof checker can formally elaborate that evidence into a focused sequent calculus. Such an elaboration thus guarantees the soundness of that proof. These papers also describe a "reference proof checker" that has been built with the expectation that its formal correctness can be established. That checker is also able to do bounded *proof reconstruction* as well as allow both deterministic and non-deterministic computation to be mixed with deduction.

Our understanding of focused proofs in the presence of both induction and co-induction (inference rules found in model checkers and most theorem provers) is less well developed. As a result, Miller and Tiu have been studied a simple approach of proof certificate in the setting of model checking in the hope of identifying the relevant proof theory designs that need to be developed. In [33], they showed how tabled deduction in model checking can be used to provide a formal proof certificate for a range of co-inductively defined predicates.

6.5. Combinations of classical and intuitionistic logic

Participant: Dale Miller.

Chuck Liang and Miller have been studying the question of how one can mix intuitionistic and classical logic into a single logic. The initial motivation for considering this problem arose from the concerns raised by the ProofCert project of how best to deal with both classical and intuitionistic logic and their associated proof evidence. Will there need to be two different kinds of checkers and two different kinds of libraries for these two different kinds of logics? Will we be able to mix theorems and proofs in one logic with those in the second logic in rich and useful ways?

One way we have considered answering this question is to actually consider a third logic that combines these other two logics. Our work on such combinations is reported in [16], where a thorough analysis of the semantics and proof theory of such a combination is provides, and in [24], where significant examples of the computational aspects of proofs are explored in detail.

6.6. Formal meta theory of sequent calculus

Participant: Dale Miller.

Keeping with the ProofCert theme of finding global, eternal, and formal mechanisms representing proof evidence, Miller and Pimentel describe in [17] a way in which linear logic can be used to formally specify inference rules for a wide range of proof system in several logics. They were able to show that adequacy of their encodings and to provide sufficient conditions for both cut-elimination and initial-elimination to hold for the resulting proof systems. The fact that these elimination results hold or not is an important characteristic for judging a proof system. Using this work, these important questions can be resulted automatically for a wide range of such proof systems.

6.7. The correctness of program using finite precision

Participants: Ivan Gazeau, Dale Miller.

Programs dealing with real number quantities must live with the fact that such numbers are represented using only finite precision. As such, programs that might be considered correct over the abstract field of infinite precision arithmetic can display chaotic and incorrect behaviors when run on actual computer hardware.

One such problem with finite precision is that programs can "leak" information about values that are intended to be hidden or at least obfuscated as happens in the area differential privacy. In [22], Gazeau, Miller, and Palamidessi illustrated just how such attacks on information hiding can be made and how it is possible to add noise to reported data values in such a way that only appropriate amounts of information leakage occurs.

In his PhD thesis, *Safe Programming in finite precision: Controlling the errors and information leaks* (École Polytechnique, 2013 [11]), Gazeau develops that theme further as well as shows how techniques from rewriting theory can be applied to show that, in some situations, the chaotic behavior of finite precision programs can be expected to converge in acceptable time to acceptable answers.

6.8. Sequent Calculus with Calls to a Decision Procedure

Participants: Mahfuza Farooque, Stéphane Graham-Lengrand.

In the PSI project, a version of the focused sequent calculus (for first-order classical logic) has been designed, which can call external decision procedures. Several results were achieved in 2013 since the last Activity Report:

Firstly, a bug was discovered in the proof of cut-elimination, which was used to prove the logical completeness of the calculus. Fixing the bug required minor changes in the definition of the system, but incurred a major re-development of the meta-theory. Out of this technical work, one idea emerged: in presence of a non-trivial theory, changing the polarity of literals may change the provability of formulas. This was quite unexpected, but it led to interesting issues, such as finding sufficient conditions on polarities to guarantee cut-elimination and logical completeness. An substantial achievement in this research topic was to successfully address such issues, which gave rise to a new version of the report [30].

Secondly, more techniques from automated reasoning were captured as proof-search in this sequent calculus (the incremental construction of proof-trees): besides the SMT-solving algorithm DPLL(T) treated successfully in 2012 (which was written down and published this year in [21]), the techniques of *clause tableaux* and *connection tableaux* were captured this year. This includes in particular a notion of *clause tableaux modulo theories* that C. Tinelli introduced in 2007 [60]. This new range of captured techniques is interesting as clause tableaux are designed to handle quantifiers, which DPLL(T) does not. This gives a new hope to combine the efficiency of SAT-solvers for propositional reasoning with the handling of quantifiers.

6.9. Path Functors in the Category of Small Categories

Participant: François Lamarche.

In [31] François Lamarche gives a detailed description of two path functors in the category of small categories, which he calls $\mathbf{Pe}$ and $\mathbf{P}$, and proves some of their important properties. The second of these is the functor which is used to model the Martin-Löf identity type in [47]; it associates to every small category X an internal category structure whose object of objects is X; one important theorem which is proved in [31] is that the category of internal (co- or contravariant) presheaves on $\mathbf{P}X$ coincides with the category of Grothendieck bifibrations over the base X. Thus, through a trivial use of monadic abstract nonsense, we can say that $\mathbf{P}X$ is the free bifribration over X. The category $\mathbf{P}X$ is obtained by taking the bigger $\mathbf{Pe}X$, which is a little more than just a category, being poset-enriched, and getting rid of the order enrichment by quotienting. $\mathbf{Pe}X$ is a more general kind of bifibration than an ordinary Grothendieck bifibration, and the enrichment is necessary to describe its properties, thus taking us outside of the theory 1-categores.

6.10. Subformula Linking as an Interaction Method

Participant: Kaustuv Chaudhuri.

We showed how to generalize the *calculus of structures*, a *deep inference* formalism, for classical linear logic to a *calculus of linking* [18]. This generalization simplifies the calculus by eliminating most of its inference rules. In its place we add a notion of annotation with *links* and a *link resolution* procedure. We show that this is sound and complete with respect to the usual calculus of structures. The linking calculus is the foundational basis of the *Profound* tool described in 5.1 .

6.11. Recovering Proof Structures in the Sequent Calculus

Participants: Kaustuv Chaudhuri, Stefan Hetzl, Dale Miller.

The *sequent calculus* is often criticized as a proof syntax because it contains a lot of noise. It records the precise minute sequence of operations that was used to construct a proof, even when the order of some proof steps in the sequence is irrelevant and when some of the steps are unnecessary or involve detours. These features lead to several technical problems: for example, cut-elimination in the classical sequent calculus LK, as originally developed by Gentzen, is not confluent, and hence proof composition in LK is not associative. Many people choose to discard the sequent calculus when attempting to design a better proof syntax with the desired properties.

In recent years, there has been a project at Parsifal to recover some of these alternative proof syntaxes by imposing a certain abstraction over sequent proofs. Our technique, pioneered at Parsifal, involves the use of *maximal multi-focusing* which gives a syntactic characterization of those sequent proofs that: (1) have a "don't care" ordering of proof steps where the order does not matter, and (2) groups larger logical steps, called *actions*, into a maximally parallel form where only important orderings of actions are recorded. The earliest example of this technique was in [40], where we showed a class of sequent proofs that were isomorphic to proof nets for multiplicative linear logic. In 2012, we were able to obtain a similar result for first-order classical logic, wherein we defined a class of sequent proofs that are isomorphic to expansion proofs, a generalization of Herbrand disjunctions that is in some sense a minimalistic notion of proof for classical logic. This result was published in a preliminary form at the CSL 2012 conference [39].

In 2013 we published an extended paper on this result in the Journal of Logic and Computation [14]. The major contribution here was a detailed proof of the result that gives a precise account of the proof identifications made by expansion proofs.

<p style="text-align:center; color:red">POEMS Project-Team</p>

6. New Results

6.1. Time domain wave propagation problems

6.1.1. *Numerical methods in electromagnetism*
Participant: Gary Cohen.

In the framework of contract GREAT, we implemented and compared two discontinuous Galerkin methods to solve Maxwell's equations for time dependent problems, the first using tetrahedral meshes (first used by Hesthaven), the second using hexahedral meshes with mass lumping. This comparison showed the undeniable superiority of the second method, 4-7 times faster (for orders from 2 to 4) for the same accuracy.

The ultimate goal of this program was the hybridization of those two types of meshes because the construction of purely hexahedral mesh for complex geometries is often difficult or almost impossible. A first approach was studied in the thesis Morgane Bergot, where the transition between the two grids was performed by the use of pyramids. The implementation of such elements is difficult and costly, we were interested in a transition mortars elements capable of hybridizing directly tetrahedra with flat faces and hexahedral with non-planar faces. This approach is promising and should lead to a rapid and efficient method. A theoretical study of the error and stability is conducted in collaboration with Eric Chung of CUHK (Chinese University of Hong Kong).

Moreover, always with E. Chung, we became interested in the construction of a discontinuous Galerkin method on hexahedral meshes offset for solving Maxwell's equations. This approach has two advantages : firstly, the shift naturally removes the spurious waves which appear with other approaches (which usually requires the introduction of a dissipative term to remove them). On the other hand, a phenomenon of super-convergence appears which should lead to a substantial time saving. A first study of the dispersion of this method led to a publication.

6.1.2. *Solving the Homogeneous Isotropic Linear Elastodynamics Equations Using Potentials and Finite Elements.*
Participants: Aliénor Burel, Patrick Joly.

The aim of this subject, investigated in collaboration with Marc Duruflé (Inria Bordeaux) and Sébastien Imperiale (Inria Saclay), is to use the classical theoretical decomposition of the elastodynamic displacement into two potentials referring to the pressure wave and the shear wave, and use it in a numerical framework. During the past two years, a method has been proposed for solving the Dirichlet problem (clamped boundary), successfully analysed and implemented, and for the free boundary conditions, we proposed an original method considering these boundary conditions as a perturbation of the Dirichlet conditions. This approach performs successfully in the time-harmonic regime but appears to give rise to severe instabilities in the time-dependent case after space and time discretization. Our investigations seem to prove that this instability is already present in the semi-discrete problem in space, but we are still looking for an explanation of this phenomenon.

6.1.3. *Limiting amplitude principle in a two-layered medium composed of a dielectric and a metamaterial*
Participants: Maxence Cassier, Christophe Hazard, Patrick Joly, Valentin Vinoles.

We are investigating this problem from both theoretical and numerical points of view. This is also the object of a collaboration with B. Gralak from the Institut Fresnel in Marseille.

This work is the time-domain counterpart of the research done at Poems about frequency domain analysis of metamaterials in electromagnetism, in the framework of the ANR Project METAMATH. One fundamental question is the link between the evolution / time harmonic problems via the limiting amplitude principle, in particular in the cases where the time harmonic problem fails to be well posed. This occurs, at certain frequencies, when one considers a transmission problem between a standard dielectric material and a dispersive material obeying for instance to the Drude model (other models as Lorentz materials or their generalization also give rise to the same results). Indeed, for well-chosen coefficients (which we refer as the critical case), there exist critical frequencies (only one frequency ω_c for the Drude model) for which the metamaterial behaves as a material whose equivalent electric permittivity and magnetic permeability are negative and precisely opposite to the ones of the dielectric medium : in such a situation, in the case of a plane interface, it is known that the time harmonic transmission problem is strongly ill-posed.

We have considered the evolution problem in a two-layered medium, when we consider a source term $f(x) \, e^{i \, \omega t}$ with frequency $\omega > 0$. In the non critical case, the limiting amplitude principle holds : for large times, the solution $u(x,t)$ of the evolution problem "converges" to a time harmonic solution of the form $u^\infty(x) \, e^{i \, \omega t}$. In the critical case, the limiting amplitude principle no longer holds. If $\omega \neq \omega_c$, the solution of the evolution problem behaves when $t \to +\infty$ to a "double frequency" solution of the form

$$u^\infty(x) \, e^{i \, \omega t} + u^\infty_{c,+}(x) \, e^{i \, \omega_c t} + u^\infty_{c,-}(x) \, e^{-i \, \omega_c t}.$$

If $\omega = \omega_c$, the solution blows up linearly at infinity :

$$u(x,t) \sim t \, u^\infty_c(x) \, e^{i \, \omega_c t} \quad (t \to +\infty)$$

where the function $u^\infty_c(x)$ is "concentrated" near the interface : this can be interpreted as an "interface resonance" phenomenon. We have performed various numerical experiments (using in particular the stabilized PMLs evoked in section 6.3.2) that illustrate this resonance phenomenon (cf figure 1).

Figure 1. Left : the interface wave. Right : solution at one point as a function of time.

From the mathematical point of view, the method we have used consists in rewriting the original problem as an abstract Schrödinger equation

$$i \, \frac{du}{dt} + Au = F \, e^{i \, \omega t}$$

where A is a self-adjoint operator in an appropriate Hilbert space H. The key of the analysis is the spectral theory of the operator A. This permits a quasi-explicit representation of the solution via the (generalized) diagonalization of A. This is achieved by combining a partial Fourier transform along the interface with Sturm-Liouville type techniques in the orthogonal direction. In the critical case, the resonance phenomenon appears to be linked to the fact that A admits a (single) eigenvalue of infinite multiplicity.

6.1.4. *Finite differences method for nonlinear acoustic waves with fractional derivatives*
Participant: Jean-François Mercier.

This subject is developed in collaboration with Bruno Lombard from LMA.

We develop a numerical method to study the wave propagation in a 1-D guide with an array of Helmholtz resonators, considering large amplitude waves and viscous boundary layers. The model consists in two coupled equations: a nonlinear PDE for the velocity in the tube (Burgers like equation) and a linear ODE describing the pressure oscillations in the Helmholtz resonators. The dissipative and dispersive effects in the tube and in the necks of the resonators are modelled by fractional derivatives expressed as convolution products with singular kernels. Based on a diffusive representation, the convolution kernels of the fractional derivatives are replaced by a finite number of memory variables that satisfy local ordinary differential equations. The procedure to compute weights and nodes of the diffusive representation of fractional derivatives is optimized. Moreover an adequate coupling between the PDE and the ODE is introduced to be sure that the discrete energy is decreasing. A splitting strategy is then applied to the evolution equations to obtain a stable scheme under the optimal CFL condition: the propagative part is solved by a standard TVD scheme for hyperbolic equations, whereas the diffusive part is solved exactly. This approach is validated by comparisons with exact solutions. The properties of the full nonlinear solutions are investigated numerically. In particular, the existence of acoustic solitary waves, due to the competition between dispersion and nonlinear effects, is confirmed.

6.2. Time-harmonic diffraction problems

6.2.1. *Numerical computation of variational integral equation methods*
Participants: Marc Lenoir, Nicolas Salles.

The discretization of 3-D scattering problems by variational boundary element methods leads to the evaluation of such elementary integrals as

$$\int_{S \times T} G\left(x, y\right) v\left(x\right) w\left(y\right) \, dx \, dy \ \text{and} \ \int_{S \times T} \frac{\partial}{\partial n_y} G\left(x, y\right) v\left(x\right) w\left(y\right) \, dx \, dy \tag{6}$$

where v and w are polynomial basis functions, G is the Green kernel and S and T two planar polygons from the discretization of the boundary. Due to the singularity of the kernel, the numerical evaluation of these integrals may lead to inaccurate results when S and T are close to each other. We split G and its gradient into a regular part which involves classical numerical techniques and a singular part subject to our method. This new method consists in integrating exactly integrals such as

$$I = \int_{S \times T} v(x) \frac{1}{\|x - y\|} w(y) \, dx \, dy \ \text{and} \ J_\zeta = \int_{S \times T} \frac{x - y}{\|x - y\|^{1+\zeta}} \, dx \, dy, \, \zeta \in \{0, 2\}. \tag{7}$$

or numerically integrals such as:

$$\mathcal{L} = \int_{S \times T} v(x) \frac{e^{ik\|x-y\|}}{\|x-y\|} w(y) \, dx \, dy \tag{8}$$

where v and w are basis functions of order 0 or 1. The general approach relies on two steps.

Basic formulas : let $f(x,d) : \Omega \subset \mathbb{R}^n \times \mathbb{R} \longrightarrow \mathbb{R}$ a positively homogeneous function of degree q. By Euler's formula and Green's theorem we have the function $I(d)$ satisfies :

$$(q+n)I(d) = dI'(d) + \int_{\partial\Omega} (\overrightarrow{z}\,|\overrightarrow{\nu}\,) f(z,d) \, d\gamma_z, \quad \text{with } I(d) = \int_\Omega f(z,d) \, dz \tag{9}$$

where $\overrightarrow{\nu}$ is the exterior normal to Ω. Provided $d^{-(q+n)} \int_\Omega f(z,d) \, dz \to 0$ as $d \to +\infty$ one obtains

$$I(d) = d^{q+n} \int_{\partial\Omega} (\overrightarrow{z}\,|\overrightarrow{\nu}\,) \int_d^{+\infty} \frac{f(z,t)}{t^{q+n+1}} dt d\gamma_z. \tag{10}$$

When $f(z,d)$ does not depend on d and $q + n \neq 0$ then

$$I = \frac{1}{q+n} \int_{\partial\Omega} (\overrightarrow{z}\,|\overrightarrow{\nu}\,) f(z) \, d\gamma_z. \tag{11}$$

As long as the inner integral in (5) can be explicitly evaluated, both formulas reduce an n-dimensional integral to an $(n-1)$ one. When Ω is an n-dimensional polyhedron (such as $S \times T$ with $n = 4$), $(\overrightarrow{z}\,|\overrightarrow{\nu}\,)$ is constant on each $(n-1)$-face of Ω, a simplification of crucial importance in the sequel.

The reduction process : we have obtained formulas for three types of geometrical configurations: S and T are (i) coplanar, (ii) in secant planes and (iii) in parallel planes. All these cases are treated using formulas (6) or (5) or both, depending on the relative positions of S and T. As an example, we present the simple but significant result for the self-influence coefficient ($S = T$). Let A_i be a vertex of the triangle, α_i the opposite side, B_i the orthogonal projection of A_i on α_i and $\gamma_i = \|A_i B_i\|$. After 3 successive reductions using formula (6), one obtains

$$I = \int_{S \times S} \frac{1}{\|x-y\|} dx dy = \frac{2|S|}{3} \sum_{i=1,3} \gamma_i R(A_i, \alpha_i), \tag{12}$$

where $R(A_i, \alpha_i)$ is given analytically by $((i, jk)$ being a circular permutation of $(1,2,3))$

$$R(A_i, \alpha_i) = \int_{\alpha_i} \frac{1}{\|A_i - y\|} dy = \arg\sinh \frac{\|B_i A_k\|}{\gamma_i} - \arg\sinh \frac{\|B_i A_j\|}{\gamma_i}. \tag{13}$$

Results for the 3-D Helmholtz equation with piecewise constant density have been obtained for all pairs of panels. Integral $\mathcal{L}$ (see formula (3)) can be reduced to a linear combination of 1-D or 2-D integrals when triangles have at least one common vertex; the resulting integrals have to be evaluated numerically but the final integrands are simple and regular on the domain of integration. For example, when $T = S$ and with piecewise constant basis functions, one has:

$$\mathcal{L} = \int_{S \times S} \frac{e^{ik\|x-y\|}}{\|x-y\|} \, dx \, dy = 4|S| \sum_{i=1}^3 \gamma_i \int_{\alpha_i} f(\|A_i - y\|) \, ds_y \tag{14}$$

where $f(r) = i \dfrac{e^{ikr} - 1 - ikr + k^2 r^2/2}{k^3 r^4}$.

The extension to linear basis functions is in progress. Our method works also for 3-D Maxwell's equations with linear edge basis functions (for MFIE and EFIE). Despite some (possibly) lengthy calculations, the principle is rather straightforward and the method is quite flexible, leading to the reduction of 4-D integrals to a linear combination 1-D regular integrals which can be numerically or even explicitly evaluated. It is possible to use our method for Collocation technique, 2-D BEM and volume integral equations. A high degree of accuracy can be obtained, even in the case of nearly singular integrals. We will present the method and some results for 3-D Helmholtz equation.

6.2.2. *Integral equations for modelling eddy current non destructive testing experiments*
Participants: Marc Bonnet, Audrey Vigneron.

This work in collaboration with E. Demaldent (CEA LIST) is concerned with developing boundary element solvers for modelling eddy current non destructive testing experiments, taking into account the probe, the probed part and the surrounding air. Attention is focused in implementing Galerkin-type formulations, overcoming ill-conditioning arising in configurations involving high contrasts, and fast solvers. Among several possible integral formulations based on either Maxwell's equations or the eddy-current model, a weighted coupled formulation using a loop-tree decomposition of the trial and test spaces was found to perform adequately over the whole range of values of physical parameters typical of eddy-current NDT experiments.

6.2.3. *Elastodynamic fast multipole method for semi-infinite domains.*
Participants: Marc Bonnet, Stéphanie Chaillat.

The use of the elastodynamic half-space Green's tensor in the FM-BEM is a very promising avenue for enhancing the computational performances of 3D BEM applied to analyses arising from e.g. soil-structure interaction or seismology. This work is concerned with a formulation and computation algorithm for the elastodynamic Green's tensor for the traction-free half-space allowing its use within a Fast Multipole Boundary Element Method (FM-BEM). Due to the implicit satisfaction of the traction-free boundary condition achieved by the Green's tensor, discretization of (parts of) the free surface is no longer required. Unlike the full-space fundamental solution, the elastodynamic half-space Green's tensor cannot be expressed in terms of usual kernels such as $e^{ikr/r}$ or $1/r$. Its multipole expansion thus cannot be deduced from known expansions, and is formulated in this work using a spatial two-dimensional Fourier transform approach. The latter achieves the separation of variables which is required by the FMM. To address the critical need of an efficient quadrature for the 2D Fourier integral, whose singular and oscillatory character precludes using usual (e.g. Gaussian) rules, generalized Gaussian quadrature rules have been used instead. The latter were generated by tailoring for the present needs the methodology of Rokhlin's group. Extensive numerical tests have been conducted to demonstrate the accuracy and numerical efficiency of the proposed FMM. In particular, a complexity significantly lower than that of the non-multipole version was shown to be achieved. A full FM-BEM based on the proposed acceleration method for the half-space Green's tensor is currently under way. This treatment of the Green's tensor will be extended to other cases, e.g. layered semi-infinite media.

6.2.4. *Domain decomposition methods for time harmonic wave propagation*
Participants: Patrick Joly, Mathieu Lecouvez.

This work is motivated by a collaboration with the CEA-CESTA (B. Stupfel) through the PhD thesis of M. Lecouvez and is the object of a collaboration with F. Collino, co-advisor of the thesis with P. Joly.

We have considered first the case of the scalar Helmholtz equation for which we have developed a non overlapping iterative domain decomposition method based on the use of Robin type transmission conditions, in the spirit of previous works in the 90's by Collino, Desprès, and Joly.

The novelty of our approach consists in using new transmission conditions using some specific impedance operators in order to improve the convergence properties of the method (with respect to more standard Robin conditions). Provided that such operators have appropriate functional analytic properties, the theory shows that one achieves geometric convergence (in opposition the slow algebraic convergence obtained with standard methods). These properties prevent the use of local impedance operator, a choice that was commonly done for the quest of optimized transmission conditions (following for instance the works of Gander, Japhet, Nataf). We propose a solution that uses nonlocal integral operators using appropriate Riesz potentials, the important feature of which being their singularity at the origin. To overcome the disadvantage of dealing with completely nonlocal operators, we suggest to work with truncated kernels, involving adequate smooth cut-off function. The results we have obtained are

- A complete theoretical justification of the exponential convergence of the algorithm in the 2D case for smooth enough interfaces. The extension to 3D is in progress : the case of a spherical interface is in particular completely understood.

- An heuristic analysis of the influence of the truncation procedure (several choices are possible) on the convergence of the method, together with a (semi-analytical) search for optimal values of the parameters involved in the method to improve the convergence rate.

- The implementation of the method in 2D and an intensive campaign of numerical validation of the method that appear to provide very good performance and seem to indicate that the method is quite robust with respect to increasing frequency (which remains to be proven). Let us however mention that, not so important but unexpected phenomena, due to space discretization, have been observed and remain to be explained. The implementation in 3D, in cooperation with M. Duruflé, is in progress.

The relevant application at CESTA being electromagnetism, the extension of the method to 3D Maxwell's equations, which proposes new non trivial difficulties, has been initiated.

As the development and the theoretical understanding of these new domain decomposition methods clearly exceed the content of one single thesis, we have proposed an ANR project on this topic, in collaboration with X. Claeys (Paris VI).

6.2.5. *Time harmonic aeroacoustics*

Participant: Jean-François Mercier.

This subject is treated in collaboration with Florence Millot (Cerfacs). We are still working on the numerical simulation of the acoustic radiation and scattering in presence of a mean flow. Up to now we have considered Galbrun's equation, but for 3D configurations it requires to introduce many unknowns. Therefore we focus now on the alternative model of Goldstein's equations. When the fluid flow and the source are potential, the acoustic perturbations are also potential and the velocity potential φ satisfy a simple scalar model. For a general flow, this model is slightly modified and is called Goldstein's equations. A new vectorial unknown ξ is introduced, satisfying a transport equation coupled to the velocity potential. φ satisfies the same modified Helmholtz's equation than in the potential flow case, in which ξ plays the role of a source term. The advantage of Goldstein's formulation compared to Galbrun's model is that the vectorial unknown vanishes in the areas where the flow is potential.

For a general flow ξ can be expressed versus φ as a convolution formula along the flow streamlines. The situation is much simpler for slow flows since the convolution formula can be simplified and the link between ξ and φ becomes explicit. Then Goldstein's equations can be solved by using continuous finite element (discontinuous elements must be used in the general case). We have proved theoretically that when replacing the general convolution formula by the "slow flow" approximation, the error on the velocity potential is small, bounded by the square of the flow velocity. This has been done for a simpler case, a shear flow, for which the streamlines are just parallel lines. Numerical tests have confirmed the square law for the error.

6.2.6. Mathematical and numerical analysis of metamaterials

Participants: Patrick Joly, Anne-Sophie Bonnet-Ben Dhia, Patrick Ciarlet, Sonia Fliss, Camille Carvalho, Valentin Vinoles, Christian Stohrer.

Metamaterials are artificial composite materials having the extraordinary electromagnetic property of negative permittivity and permeability at some frequencies . Both of sign-changing coefficients and high contrast homogenization raise new mathematical and numerical challenges. The ANR METAMATH is devoted to the study of those problems. We perform analysis both in time domain (see sections 6.1.3 and 6.3.2) and harmonic domain.

6.2.6.1. Time-harmonic transmission problems involving metamaterials

A special interest is devoted to the transmission of an electromagnetic wave between two media with opposite sign dielectric and/or magnetic constants. As a matter of fact, applied mathematicians have to address challenging issues, both from the theoretical and the discretization points of view. In particular, it can happen that the problem is not well-posed in the classical frameworks (H^1 for the scalar case, $H(curl)$ for the vector case). During 2013, we addressed the issues below.

The numerical analysis of the well-posed scalar eigenproblem discretized with a classical, H^1 conforming, finite element method, for arbitrarily shaped interfaces can be carried out with the help of T-coercivity. This work complements the paper Chesnel-Ciarlet, published in Numerische Mathematik, which handled simpler interface configurations (see also §6.2.6.2).

As a second topic, we investigated the case of a scattering problem with a 2D corner interface which can be ill-posed (in the classical H^1 framework). When this is the case, the part of the solution which does not belong to H^1 can be described as a wave which takes an infinite time to reach the corner: this "black-hole" phenomenon is observed in other situations (elastic wedges for example). We have proposed a numerical approach to approximate the solution which consists in adding some Perfectly Matched Layers in the neighbourhood of the corner. As an alternate choice, a T-coercivity approach is also being currently developed to solve the discrete problem.

Last, we studied the transmission problem in a purely 3D electromagnetic setting from a theoretical point of view. We proved that the Maxwell problem is well-posed if and only if the two associated scalar problems (with Dirichlet and Neumann boundary conditions) are well-posed. Numerical analysis of the discretized models (edge elements) is under way.

L. Chesnel left our project in March 2013 after he completed his PhD thesis on these topics. He is currently a post-doc fellow at Aalto University (Finland).

6.2.6.2. Modeling of plasmonic devices

Plasmonic surface waves occur at the interface between the vacuum (or a dielectric) and a metal, at optical frequencies, when the dielectric permittivity ε of the metal has a small imaginary part and a large negative real part. Neglecting the dissipation effects, we have to study electromagnetic problems with a sign-changing ε. An in-depth analysis has been done by Lucas Chesnel during his PhD. In the context of the PhD of Camille Carvalho, we extended the results obtained previously by Lucas Chesnel to more realistic configurations. First, we studied the diffraction of a transversely polarized plane wave by a cylindrical metallic inclusion, when the section of the inclusion presents edges (cf. §6.2.6.1). Then, we considered a related spectral problem in view of studying plasmonic guided waves. The spectral theory is far from obvious. In particular, we have to introduce a non-selfadjoint formulation which provides physical real eigenvalues and complex spurious ones. For both the diffraction problem and the spectral problem, a MATLAB code has been developed, where Perfectly Matched Layers are introduced at the corners to take into account the presence of black-hole waves seemingly absorbed by the corners. The convergence of the finite element discretization (including convergence of the eigenvalues) has been proved (see §6.2.6.1).

6.2.6.3. Study of metamaterials via numerical homogenization

Recently, we have started to study the numerical approximation of the full models, using the HMM (Heterogeneous Multiscale Method). Recall that the full model is obtained via periodization of a local model that includes slow and fast variations. With this HMM approach, computations are carried out on a global mesh, whereas the action of the test-functions is computed at a local level to take into account the fast variations. As a first step, we have begun by the application of HMM for the time-harmonic scalar problem. The case of uniformly bounded coefficients has been addressed. The more general case of non-uniformly bounded coefficients, also called the high-contrast case, is now under scrutiny. It is hoped that one can recover some extra-ordinary properties of the metamaterials with this latter case.

C. Stohrer arrived as a post-doc fellow this fall.

6.3. Absorbing boundary conditions and absorbing layers

6.3.1. *New transparent boundary conditions for time harmonic acoustic and elastic problems in anisotropic media*

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

This topic is developed in collaboration with Vahan Baronian (CEA). Non destructive testing (NDT) is a common method to check the quality of structures and is widely used in industrial applications. Typically, in aircraft design, it is required to control structures like plates. Efficient and accurate numerical methods are required to simulate NDT experiments.

In our case, we want to study the diffraction of a time harmonic wave by a bounded defect in an infinite anisotropic elastic plate. The difficulty is to find a way to restrict the finite element computation to a small box containing the defect. Indeed classical methods such as the perfectly matched layers fail when the medium is anisotropic.

Up to now we considered the simpler case of an infinite dissipative 2D medium.

Our idea, inspired by the work of Sonia Fliss and Patrick Joly for periodic media, is to consider five domains recovering the whole plane:

- a square that surrounds the defect in which we have a finite element representation of the solution,
- and four half-spaces parallel to the four edges of the square, in which we can give an analytical representation of the solution thanks to the Fourier transform.

The different unknowns are coupled by well-chosen transmission relations which ensure the compatibility between the five representations.

The method has been validated successfully in the case of anisotropic acoustic media and the implementation for the case of elasticity is in progress. The mathematical properties of the formulation and the efficiency of the method strongly depend on the presence or not of overlaps between the finite element box and the four half-planes. The formulation with overlaps has good Fredholm properties but the well-posedness for all frequencies is proved only for the formulation without overlaps.

6.3.2. *Perfectly Matched Layers in negative index metamaterials*

Participants: Patrick Joly, Eliane Bécache, Valentin Vinoles.

The simulation of waves in unbounded domains requires methods to artificially truncate the computational domain. One of the most popular ones to do so is the Perfectly Matched Layers (PMLs) which are effective and stable for non dispersive isotropic media. For non dispersive anisotropic media, we established a necessary stability condition in 2004 : PMLs are unstable in presence of so called-backward waves.

We are interested here in dispersive media and more specifically in Negative Index Metamaterials (NIMs), also called left-handed media. Those media have negative permittivity and permeability at some frequencies due to microscopic resonating structures. Since the 1990s, NIMs are the subject of active researches due to their promising applications : superlens, cloaking, improved antenna, etc.

In a first step, we consider a simple model of NIMs : the Drude model. For this model, a plane wave analysis shows the simultaneous presence of both forward and backward waves and numerical simulations confirm the instability of standard PMLs (cf figure 2) that result from complex changes of variable leading to the following modification of the spatial derivatives

$$\partial_x \longrightarrow \left(1 + \frac{\sigma_x(x)}{i\omega}\right)^{-1} \partial_x \quad \text{and} \quad \partial_y \longrightarrow \left(1 + \frac{\sigma_y(y)}{i\omega}\right)^{-1} \partial_y$$

where $\sigma_x(x) > 0$ and $\sigma_y(y) > 0$ are the damping terms. Inspired by works of the physics community, we propose more general changes of variable

$$\partial_x \longrightarrow \left(1 + \frac{\sigma_x(x)}{i\omega\psi(\omega)}\right)^{-1} \partial_x \quad \text{and} \quad \partial_y \longrightarrow \left(1 + \frac{\sigma_y(y)}{i\omega\psi(\omega)}\right)^{-1} \partial_y$$

where $\psi(\omega)$ is a function to be chosen judiciously. We have generalised the previous necessary stability condition for those new PMLs, called Stabilized Perfectly Matched Layers, for dispersive media. This analysis allows us to understand the instabilities observed for standard PMLs in NIMs and to propose a choice of functions $\psi(\omega)$ which take into account the backward waves and stabilize the PMLs as confirmed by numerical simulations (cf figure 2).

Figure 2. Left : the standard PMLs are unstable. Right : the Stabilized PMLs are stable.

6.3.3. Perfectly Matched Layers in plasmas

Participants: Patrick Joly, Eliane Bécache, Valentin Vinoles.

This work was done during the internship of Guillaume Chicaud in the framework of the ANR CHROME which concerns the study of electromagnetic wave propagation in plasmas. Our aim is to develop efficient and robust codes to simulate wave propagation in unbounded plasmas models. The simulation of waves in plasmas requires technics to bound the computational domain. As plasmas are dispersive media where backward waves may occur, the difficulties to construct stable PMLs are analogous to the ones encountered for Negative Index Metamaterials (cf 6.3.2). This work is a preliminary study of this topics, in a simplified model, the case of a 2D anisotropic uniaxial plasma. It consists first in analyzing the presence of backward waves with a plane wave analysis. The second step was to implement the equations using standard PMLs and to confirm the expected

instabilities. Finally, we proposed stabilized PMLs (SPMLs), inspired by the work done in metamaterials (see section 6.3.2).

The continuation of this project will constitute the subject of the post-doc of Maryna Kachanovska.

6.4. Waveguides, resonances, and scattering theory

6.4.1. *An improved modal method in non uniform acoustic waveguides*

Participant: Jean-François Mercier.

This topic is developed in collaboration with Agnès Maurel (Langevin Institute ESPCI).
We develop modal methods to study the scattering of an acoustic wave in a non uniform waveguide. Usual modal approaches are efficient only when a rather large number of evanescent modes are taken into account. An improved representation has been proposed in which an additional transverse mode and an additional unknown modal component are introduced. This so called boundary mode helps to better satisfy the Neumann boundary conditions at the varying walls. A system of coupled ordinary differential equations is obtained and is found to remain coupled in the straight part of the waveguide which implies that the classical radiation condition cannot be applied directly at the inlet/outlet of the scattering region.

We revisit the coupled mode equations in order to derive an improved system, in which the additional mode can be identified as evanescent mode, and then adapted to define radiation conditions. This makes possible the implementation of efficient numerical multimodal methods (like the admittance matrix method) and also approximate solutions can be found using the Born approximation. The numerical tests have shown that our method is very efficient to reduce the number of degree of freedom: adding to the boundary mode, it is sufficient to take only the propagative modes to get very good results. This is in particular interesting at low frequency when only the plane mode propagates. In the low frequency regime, the system can be solved analytically, using the Born approximation, leading to improved approximate equations compared to the usual Webster's approximation.

6.4.2. *Construction of non scattering perturbations in a waveguide*

Participants: Anne-Sophie Bonnet-Ben Dhia, Eric Lunéville.

This work is done in collaboration with Sergei Nazarov from Saint-Petersbourg University and during the internship of Yves Mbeutcha. We consider a two-dimensional homogeneous acoustic waveguide and we aim at designing deformations of the boundary which are invisible at a given frequency (or more generally at a finite number of given frequencies) in the sense that they are non scattering. To find such invisible perturbations, we take advantage of the fact that there are only a finite number of propagative modes at a given frequency in a waveguide. As a consequence, the invisibility is achieved by canceling a finite number of scattering coefficients, and an invisible deformation only produces an exponentially decreasing scattered field, not measurable in the far field.

The first step consists in studying the effect of a small deformation, of amplitude ε. The asymptotic analysis allows to derive the first order terms of the scattering coefficients, as integrals involving the function describing the deformation. This leads to express the deformation as a linear combination of some explicit (compactly supported) functions, so that invisibility is satisfied if and only if the coefficients of the linear combination are solution of a fixed point equation. The key point is that we can prove, using the results of the asymptotic analysis, that the function of this fixed point equation is a contraction for ε small enough. This proves the existence of invisible deformations of amplitude ε. Moreover, it provides a natural algorithm to compute the invisible deformation.

This has been tested numerically and the results are in perfect agreement with the theory. At low frequency, the good news is that ε can be taken quite large (the amplitude of the deformation may be half the size of the guide). But this deteriorates when the frequency increases.

6.4.3. *Localized modes in unbounded perturbed periodic media*

Participants: Patrick Joly, Sonia Fliss, Elizaveta Vasilevskaya.

This topic is investigated in collaboration with Bérangère Delourme (Univ. Paris XIII) and constitutes the subject of the E. Vasilevskaya's PhD thesis. We are interested in a 2D propagation medium which is a localized perturbation of a reference homogeneous periodic reference medium. This reference medium is a "thick graph", namely a thin structure (the thinness being characterized by the parameter $\delta > 0$) whose limit when δ tends to 0 is a periodic graph. This is for instance the case of the thick periodic ladder and the thick periodic rectangular grid of figure. The perturbation consists in changing only the geometry (and not the material properties) of the reference medium by modifying the thickness of one of the lines of the reference medium as illustrated by figure with the perturbed ladder and perturbed grid (see figure 3). The question we investigate is whether such a geometrical perturbation is able to produce localized eigenmodes (for the ladder) or guided modes (for the grid). We have investigated this question when the propagation model is the scalar Helmholtz equation with Neumann boundary conditions (in opposition to Dirichlet conditions that have been more studied in the literature - see the works by S. Nazarov for instance). This amounts to solving an eigenvalue problem for the Laplace operator in an unbounded domain : the associated self-adjoint operator has a continuous spectrum with a band gap structure and the eigenvalues are searched in the gaps.
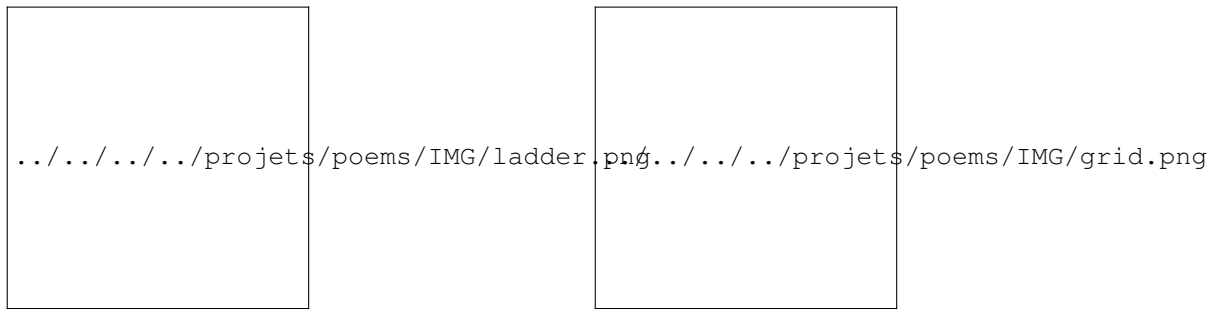

Figure 3. Left : periodic ladder (non perturbed/perturbed). Right : periodic "thick graph" (non perturbed/perturbed). The propagation domain is in grey.

With Neumann boundary conditions, we can use for the theoretical study an asymptotic analysis with respect to δ: indeed, it is well known (see in particular the works by Exner, Kuchment, Post) the limit model when δ tends to 0, is the Helmholtz equation on the graph : 1D Helmholtz equations on each branch completed by continuity and Kirchoff transmission conditions at each node. The geometrical perturbation of the original medium results into a perturbation of the Kirchoff conditions on the nodes of the modified line. The spectral analysis of the limit problem can be done completely by hand and the existence of eigenmodes for the thick medium is ensured, for δ small enough, by the existence of corresponding eigenmodes for a limit "1D operator" whose spectrum appears to possess an infinity of band gaps in each of which eigenvalues can exist, due to the perturbation. Following this idea, we have been able to prove the existence of localized modes in the case of the ladder provided that the geometrical perturbation consists in diminishing the width of one rung. One can even prove that one can produce more and more localized modes, corresponding to larger and larger frequencies, when δ is smaller and smaller. On the contrary, we conjecture that there is no localized modes when we enlarge the rung. The extension of these results to the existence of guided modes in the case of the grid in progress.

For the numerical computation of such localized modes, we have adapted the DtN approach discussed in the activity report of 2012. We gave in figure 4 an example of computed localized mode in the case of the ladder : this mode is geometrically confined at the neighbourhood of the modified rung.
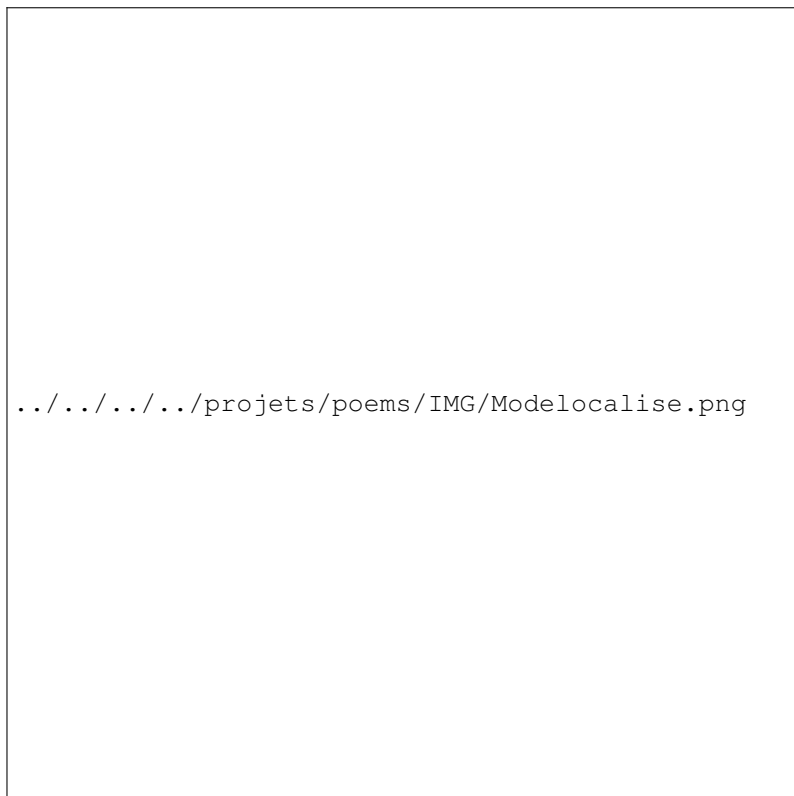
../../../../projets/poems/IMG/Modelocalise.png

Figure 4. Localized mode in the perturbed ladder.

6.5. Asymptotic methods and approximate models

6.5.1. *Homogenization and interfaces*

Participants: Sonia Fliss, Valentin Vinoles.

This topic is developed in collaboration with Xavier Claeys (LJLL, Paris VI).

The mathematical modelling of electromagnetic metamaterials and the homogenization theory are intimately related because metamaterials are precisely constructed by a periodic assembly of small resonating micro-structures involving dielectric materials presenting a high contrast with respect to a reference medium. In the framework of the ANR Metamath (see 6.2.6), we wish to look carefully at the treatment of boundaries and interfaces that are generally poorly taken into account by the first order homogenization.

This question is already relevant for standard homogenization (ie without high contrast). Indeed, the presence of a boundary induces a loss of accuracy due to the inadequateness of the standard homogenization approach to take into account boundary layer effects. Our objective is to construct approximate effective boundary conditions that would restore the desired accuracy.

We first considered a plane interface between a homogeneous and the periodic media in the standard case without high-contrast. We obtained high order transmission conditions between the homogeneous media and the periodic media. The technique we used involves matched asymptotic expansions combined with standard homogenization ansatz. Those conditions are non standard : they involve Laplace-Beltrami operators at the interface and requires to solve cell problems in infinite periodic waveguides. The derivation of the corresponding error estimates is in progress. The analysis is based on a original combination of Floquet-Bloch and a periodic version of Kondratiev technique.

The next step will be to consider the same problem but with a high-contrast periodic media in collaboration with Guy Bouchitté, a french expert in high contrast homogenization.

6.5.2. *Effective boundary conditions for thin periodic coatings*

Participants: Mathieu Chamaillard, Patrick Joly.

This topic is the object of a collaboration with Houssem Haddar (CMAP École Polytechnique). We are interested in the construction of "equivalent" boundary condition for the diffraction of waves by an obstacle with smooth boundary Γ covered with a thin coating of width δ whose physical characteristics vary "periodically" along Γ with a period proportional to the small parameter δ. For a general boundary Γ, the notion of periodicity is ambiguous: we have chosen to define the coating as the image, or the deformation, by a smooth mapping ψ_Γ of a flat layer of width δ (the reference configuration) that preserves the normals, which appears consistent with a manufacturing process. The electromagnetic parameters in the coating are then defined as the images through ψ_Γ of periodic functions in the reference configuration.

We have first considered the case of the scalar wave equation when the homogeneous Neumann condition is applied on the boundary of the obstacle. Using an asymptotic analysis in δ, which combines homogenization and matched asymptotic expansions, we have been able to establish a second order boundary condition of the form

$$\partial_\nu u + (\delta B_\Gamma^1 + \delta^2 B_\Gamma^2)u = 0,$$

where B_Γ^1 and B_Γ^2 are second order tangential differential operators along Γ. The coefficients of these operators depend on both the geometrical characteristics of Γ (through the curvature tensor), the deformation mapping ψ_Γ and the material properties of the coating, through the resolution of particular unbounded cell problems in the flat reference configuration. When the coating is homogeneous, we have checked that one recovers the well known second order thin layer condition. We have moreover proven that this approximate condition provides in $\mathcal{O}(\delta^3)$.

6.5.3. *Thin Layers in Isotropic Elastodynamics*
Participants: Marc Bonnet, Aliénor Burel, Patrick Joly.

This research is concerned with the numerical modelling of non-destructive testing experiments using ultrasonic waves. Some materials, e.g. composite materials, involve thin layers of resin. The numerical modelling of such thin layers can be problematic as they result in very small spatial mesh sizes. To alleviate this difficulty, we develop an approach based on an asymptotic analysis with respect to the layer thickness ε, aiming to model the thin layer by approximate effective transmission conditions (ETCs), which remove the need to mesh the layer. So far, ETCs that are second-order accurate in ε have been formulated, justified, implemented and numerically validated, for 2-D and 3-D configurations involving planar interfaces of constant thickness. In particular, the continuous evolution problem is shown to be stable, and a time-stepping scheme that essentially preserves the stability requirement on the time step is proposed. Extension of this work to 2-D and 3-D configurations involving a curved layer is ongoing.

6.5.4. *Mathematical modelling of electromagnetic wave propagation in electric networks.*
Participants: Geoffrey Beck, Patrick Joly.

This topic is developed in collaboration with S. Imperiale (Inria Saclay) in the framework of the ANR project SODDA, in collaboration with CEA-LETI, about the non destructive testing of electric networks. This is the subject of the PhD thesis of G. Beck.

We investigate the question of the electromagnetic propagation in thin electric cables from a mathematical point of view via an asymptotic analysis with respect to the (small) transverse dimension of the cable: as it has been done in the past in mechanics for the beam theory from 3D elasticity, we use such an approach for deriving simplified effective 1D models from 3D Maxwell's equations.

During last year, we have achieved some progress in various directions:

- Single wire coaxial cables. This is the direct continuation of what has been done last year. Concerning the lowest order, the telegraphist's model, we have extended the error analysis, previously restricted to non lossy cylindrical cables to very general cases. Technically, this relies on time Laplace transform and new, parameter dependent, Poincaré-Friedrichs inequalities. From the numerical point of view, in collaboration with M. Duruflé, we have initiated a quantitative comparison between the full 3D model and our 1D model. Furthermore we have derived and studied a higher order generalized telegraphist's equation that include dispersive effects through nonlocal capacity and inductance operators. The corresponding mathematical analysis is in progress.

- Multiple wires cables. The objective here was to extend our approach to cables containing N conducting wires. Our results into a vectorial generalized telegraphist's model with $2N$ (2 for each wire) 1D unknowns, N electrical potentials and N currents. This model involves in particular a capacity matrix C, an inductance matrix L, a resistance matrix R and a conductance matrix G, whose properties have been deeply investigated, which allowed us to justify rigorously and extend some results from the electrical engineering literature. In the most general case, the effective models also involve time memory terms with matrix valued convolution kernels.

- Junction of cables. This is a new and essential step towards the modelling of networks. We have started the case of junctions of single wire cables via the method of matched asymptotic expansions in the spirit of the PhD thesis of A. Semin.

6.5.5. *Elastic wave propagation in strongly heterogeneous media*
Participants: Simon Marmorat, Patrick Joly.

This subject enters our long term collaboration with CEA-LIST on the development of numerical methods for time-domain non destructive testing experiments using ultra-sounds, and is realized in collaboration with Xavier Claeys (LJLL, Paris VI). We aim at developing an efficient numerical approach to simulate the propagation of waves in a medium made of many small heterogeneities, embedded in a smooth (or piecewise smooth) background medium, without any particular assumption (such as periodicity) on the spatial distribution of these heterogeneities. The figure 5 is a snapshot of a simulation inside such a medium, computed thanks to classical simulation tools: to reach satisfying accuracy, one has to use mesh refinement in the vicinity of the heterogeneities, which greatly increases the computational cost of the method.

../../../../projets/poems/IMG/snap2.png

Figure 5. Snapshot of a simulation in the medium of interest, using high-order finite element method as well as local mesh refinement and local time stepping around the heterogeneities.

By considering the medium with defects as a perturbation of the smooth one, we have derived an auxiliary model in the acoustic case, involving the defect-free wave operator and some volume Lagrange multipliers which account for the presence of the defects. These Lagrange multipliers are unknown functions defined on the defects and live in some infinite dimensional functional space. Exploiting the smallness of the defects, we have shown thanks to matched asymptotic analysis that the aforementioned functional space may be well described by a finite number N of profile functions: we propose an asymptotic model by looking for the Lagrange multipliers into the space spanned by these N profile functions, and we have shown that the error hence made is controlled by ε^N, ε being the characteristic size of the defects, assumed to be small.

On a computational point of view, the asymptotic model is much easier to solve than the original one since it can be discretized using a computation mesh that ignores the presence of the heterogeneities, the Lagrangian multipliers being computed by solving a linear system of size N. A resolution of this model has been implemented in the 1D and in the 2D case, and a rigorous error estimate has been established.

6.6. Imaging and inverse problems

6.6.1. *Sampling methods in waveguides*

Participants: Laurent Bourgeois, Sonia Fliss, Eric Lunéville, Anne-Claire Egloffe.

First, we have adapted the modal formulation of sampling methods (Linear Sampling Method and Factorization Method) to the case of a periodic waveguide in the acoustic case. This study is based on the analysis of the far field of scattering solutions in cylindrical waveguides, in particular for the fundamental solution, which enables us to obtain a far field formulation of sampling methods, and then a modal formulation of such methods. The aim of the inverse problem is to retrieve a defect (that is a loss of periodicity) from the scattered fields which correspond to the incident fields formed by the Floquet modes. Some convincing numerical experiments have shown the feasibility of the method. Secondly, going back to the homogeneous waveguide in the acoustic case, we have started a study of the sampling methods in a more realistic situation, that is the data (emission and reception) are measured on the boundary of the waveguide in the time domain. This was the subject of Anne-Claire Egloffe's post-doc. The aim is to use the modal formulation of the sampling methods at all frequencies and recompose the best possible image of the defect. Some first encouraging results have been obtained when the spectrum of the incident signal is centered at a rather low frequency (corresponding to 3 propagating guided modes).

6.6.2. *Space-time focusing on unknown scatterers*

Participants: Maxence Cassier, Patrick Joly, Christophe Hazard.

This topic concerns the studies about time-reversal in the context of Maxence Cassier's thesis. We are motivated by the following challenging question: in a propagative medium which contains several unknown scatterers, how can one generate a wave that focuses selectively on one scatterer not only in space, but also in time? In other words, we look for a wave that 'hits hard at the right spot'. Such focusing properties have been studied in the frequency domain in the context of the DORT method ("Decomposition of the Time Reversal Operator"). In short, an array of transducers first emits an incident wave which propagates in the medium. This wave interacts with the scatterers and the transducers measure the scattered field. The DORT method consists in doing a Singular Value Decomposition (SVD) of the scattering operator, that is, the operator which maps the input signals sent to the transducers to the measure of the scattered wave. It is now well understood that for small and distant enough scatterers, each singular vector associated with a non zero singular value generates a wave which focuses selectively on one scatterer. Can we take advantage of these spatial focusing properties in the frequency domain to find the input signals which generate a time-dependent wave which would be also focused in time? Since any frequency superposition of a family of singular vectors associated with a given scatterer leads to a spatial focusing, the main question is to synchronize them by a proper choice of their phases. The method we propose is based on a particular SVD of the scattering operator related to its symmetry. The signals we obtain do not require the knowledge of the locations of the scatterers. We compare it with some "optimal" signals which require this knowledge. Our study is illustrated by a two dimensional acoustic model where both scatterers and transducers are assumed pointlike (see figure 6).

6.6.3. *The exterior approach to retrieve obstacles*

Participant: Laurent Bourgeois.

This theme is a collaboration with Jérémi Dardé from IMT (Toulouse). The aim is to find a fixed Dirichlet obstacle in a bounded domain by using some redundant boundary conditions (Cauchy data) on the accessible part of the boundary, while the boundary conditions are unknown on the inaccessible part of the boundary. We wish to adapt the exterior approach developed for the Laplace equation and the Stokes system to the case of time evolution problems, in particular the heat equation. The exterior approach consists in defining a decreasing sequence of domains that converge in some sense to the obstacle. More precisely, such iterative approach is based on a combination of a quasi-reversibility method to update the solution of the ill-posed Cauchy problem outside the obstacle obtained at previous iteration and of a level set method to update the obstacle with the help of the solution obtained at previous iteration. We have already introduced two different mixed formulations of quasi-reversibility for the ill-posed heat equation with lateral Cauchy data in order to use standard Lagrange finite elements.

../../../../projets/poems/IMG/figure5projets/poems/IMG/figure5projets/poems/IMG/f

../../../../projets/poems/IMG/figure5projets/poems/IMG/figure5projets/poems/IMG/f

Figure 6. The case of a scattering reference medium perturbed by two obstacles (the white circles) : modulus of the field generated by 128 transducers (left edge of each figure) at different times.

6.6.4. *Uniqueness and stability of inverse problems*

Participant: Laurent Bourgeois.

In collaboration with Laurent Baratchart and Juliette Leblond from APICS (Nice), we have proved uniqueness for the inverse Robin problem with a boundary coefficient in L^∞ in the 2D case, for the Laplace equation in the divergence form. The result is based on complex analysis. We have also established an abstract Lipschitz stability result for inverse problems such that the set of parameters is a compact and convex subset of a finite dimensional space. In particular, such result can be applied to the previous inverse Robin problem.

6.6.5. *Interior transmission problem*

Participant: Anne-Sophie Bonnet-Ben Dhia.

This work is in collaboration with Lucas Chesnel (Aalto University, Finland). During this year, we investigated a two-dimensional interior transmission eigenvalue problem for an inclusion made of a composite material. This problem plays a central role in the theory of the corresponding inverse problem. We considered configurations where the difference between the parameters of the composite material and the ones of the background change sign on the boundary of the inclusion. In a first step, under some assumptions on the parameters, we extended the variational approach of the T-coercivity to prove that the transmission eigenvalues form at most a discrete set. In the process, we also provided localization results. Then, we study what happens when these assumptions are not satisfied. The main idea is that, due to very strong singularities that can occur at the boundary, the problem may lose Fredholmness in the natural H^1 framework. Using Kondratiev theory, we proposed a new functional framework where the Fredholm property is restored.

6.6.6. *Flaw identification using elastodynamic topological derivative or transmission eigenvalues*

Participants: Marc Bonnet, Rémi Cornaggia.

This work is in collaboration with C. Bellis (LMA, CNRS, Marseille), F. Cakoni (Univ. of Delaware, USA) and B. Guzina (Univ. of Minnesota, USA). The concept of topological derivative (TD) quantifies the perturbation induced to a given cost functional by the nucleation of an infinitesimal flaw in a reference defect-free body, and may serve as a flaw indicator function. In this work, the TD is derived for three-dimensional crack identification exploiting over-determined transient elastodynamic boundary data. This entails in particular the derivation of the relevant polarization tensor, here given for infinitesimal trial cracks in homogeneous or bi-material elastic bodies. Simple and efficient adjoint-state based formulations are used for computational efficiency, allowing to compute the TD field for arbitrarily shaped elastic solids. The latter is then used as an indicator function for the spatial location of the sought crack(s). The heuristic underpinning TD-based identification, which consists in deeming regions where the TD is most negative as the likeliest locations of actual flaws and on formulating higher-order topological expansions in the elastodynamic case, has (with C. Bellis and F. Cakoni) been given a partial justification within the limited framework of acoustic inverse scattering using far-field data. Current investigations (M. Bonnet, R. Cornaggia) include setting up and justifying the formulation of higher-order topological expansions for the elastostatic and elastodynamic cases.

Another ongoing research on a related topic addresses the use of transmission eigenvalues (TEs), i.e. values of the wave number for which the homogeneous interior transmission problem (ITP) related to the scattering scattering of time-harmonic elastic waves by aninhomogeneity D admits non-trivial solutions. This works (R. Cornaggia, in collaboration with C. Bellis, F. Cakoni, B. Guzina) aims on the one hand to understand better how to compute the TEs -if any- in the case where D's characteristics vary periodically. On the other hand it looks for how a previously obtained knowledge of the TE set could be the basis of an identification process. In a preliminary study considering 1-D elastic beams with periodically varying section over a length L, gradient elasticity was found to be a well-suited homogenization model to both compute the TEs and identify L, the periodic cell length and the damage parameter from available values of the TEs.

6.6.7. *Topological derivative in anisotropic elasticity*
Participant: Marc Bonnet.

This work is in collaboration with G. Delgado (PhD student, CMAP Ecole Polytechnique and EADS IW). Following up on previous work on the topological derivative (TD) of displacement-based cost functionals in anisotropic elasticity, a TD formula has been derived and justified for general cost functionals that involve strains (or displacement gradients) rather than displacements. The small-inclusion asymptotics of such cost functionals are quite different than in the previous case, due to the fact that the strain perturbation inside an elastic inclusion has a finite, nonzero asymptotic value in the limit of a vanishingly small inclusion. Cost functionals of practical interest having this format include von Mises equivalent stress (often used in plasticity or failure criteria) and energy-norm error functionals for coefficient-identification inverse problems. This TD formulation has been tested on 2D and 3D numerical examples, some of them involving anisotropic elasticity and nonquadratic cost functionals.

6.6.8. *Energy functionals for elastic medium reconstruction using transient data*
Participant: Marc Bonnet.

This work is in collaboration with W. Aquino (Duke Univ., USA). Energy-based misfit cost functionals, known in mechanics as error in constitutive relation (ECR) functionals, are known since a long time to be well suited to (electrostatic, elastic,...) medium reconstruction. In this ongoing work, a transient elastodynamic version of this methodology is developed, with emphasis on its applicability to large time-domain finite element modeling of the forward problem. The formulation involves coupled transient forward and adjoint solutions, which greatly hinders large-scale computations. A computational approach combining an iterative treatment of the coupled problem and the adjoint to the discrete Newmark time-stepping scheme is found to perform well on cases where both the FE model and the identification problem are of large size (2D and 3D elastodynamic numerical experiments made so far involve up to half a million unknown for the discretized inverse problem), making the time-domain ECR functional a worthwhile tool for medium identification.

<p align="center" style="color:red">Popix Team</p>

6. New Results

6.1. Estimation in mixed-effects diffusion models

Participant: Marc Lavielle.

We have coupled the SAEM algorithm and the extended Kalman filter for maximum likelihood estimation in mixed-effects diffusion models: we have considered some general mixed-effects diffusion models, in which observations are made at discrete time points and include measurement errors. In these models, the observed likelihood is generally not explicit, making maximum likelihood estimation of the parameters particularly complex. We have proposed a specific inference methodology for these models. In particular, it combines the SAEM algorithm with the extended Kalman filter to estimate the population parameters. We have also provided some tools for estimating the individual parameters, for recovering the individual underlying diffusion trajectories and for evaluating the model. We evaluated the methods on simulations and applied them to a pharmacokinetics example.

6.2. Estimation in mixtures of models

Participant: Marc Lavielle.

We have proposed an improved SAEM algorithm for maximum likelihood estimation in mixtures of non linear mixed effects models. This involves a new methodology for maximum likelihood estimation in mixtures of non linear mixed effects models (NLMEM). Such mixtures of models include mixtures of distributions, mixtures of structural models and mixtures of residual error models. Since the individual parameters inside the NLMEM are not observed, we have proposed to combine the EM algorithm usually used for mixtures models when the mixture structure concerns an observed variable, with the Stochastic Approximation EM (SAEM) algorithm, which is known to be suitable for maximum likelihood estimation in NLMEM and also has nice theoretical properties. The main advantage of this hybrid procedure is to avoid a simulation step of unknown group labels required by a "full" version of SAEM. The resulting MSAEM (Mixture SAEM) algorithm is now implemented in the MONOLIX software. We have also proposed several criteria for classification of subjects and estimation of individual parameters. Our numerical experiments on simulated data have shown that MSAEM performs well in a general framework of mixtures of NLMEM. Indeed, MSAEM provides an estimator close to the maximum likelihood estimator in very few iterations and is robust with regards to initialization. Our application of the method to pharmacokinetic (PK) data demonstrated the potential of the method for practical applications.

6.3. Moving meshes with freefem++

Participants: Astrid Decoene, Bertrand Maury.

The Arbitrary Lagrangian-Eulerian framework allows to compute free surface flows with the Finite Element functions defined on a fittedmesh which follows the globalmotion of the fluid domain. We have described how freefem++ can be used to implement this method, and we have provided two and three dimensional illustrations in the context of water waves.

6.4. Modeling of the oxygen transfer in the respiratory process

Participant: Bertrand Maury.

We have proposed an integrated model for oxygen transfer into the blood, coupled with a lumped mechanical model for the ventilation process. We aim at investigating oxygen transfer into the blood at rest or exercise. The first task consists in describing nonlinear effects of the oxygen transfer under normal conditions. We also include the possible diffusion limitation in oxygen transfer observed in extreme regimes involving parameters such as alveolar and venous blood oxygen partial pressures, capillary volume, diffusing capacity of the membrane, oxygen binding by hemoglobin and transit time of the red blood cells in the capillaries. The second task consists in discussing the oxygen concentration heterogeneity along the path length in the acinus

6.5. Congestion-driven dendritic growth

Participant: Bertrand Maury.

In order to observe growth phenomena in biology where dendritic shapes appear, we have proposed a simple model where a given population evolves feeded by a diffusing nutriment, but is subject to a density constraint. The particles (e.g., cells) of the population spontaneously stay passive at rest, and only move in order to satisfy some constraint, by choosing the minimal correction velocity so as to prevent overcongestion. We treat this constraint by means of projections in the space of densities endowed with the Wasserstein distance, defined through optimal transport. This allows to provide an existence result and suggests some numerical computations, in the same spirit of what the authors did for crowd motion (but with extra difficulties, essentially due to the fact that the total mass may increase). The numerical simulations show, according to the values of the parameter and in particular of the diffusion coefficient of the nutriment, the formation of dendritic patterns in the space occupied by cells.

<p align="center">**REGULARITY Project-Team**</p>

6. New Results

6.1. Stochastic integration with respect to the Rosenblatt process.

Participant: Benjamin Arras.

From a theoretical perspective to more concrete applications, fractional Brownian motion (fbm) is a fruitful and rich mathematical object. From its stochastic analysis, initiated during the nineties, several theories of stochastic integration have emerged so far. Indeed, fbm is, in general, not a semimartingale neither a Markov process. These theories rely on different properties of the stochastic integrator process and are then of different natures. Despite the quite large number of these strategies, we can group them into two fundamentally distinct categories: the pathwise and the probabilistic approaches. The probabilistic one requires highly evolved stochastic analysis tools. Indeed, the Malliavin calculus as well as Hida's distribution theory have been used in order to define stochastic integration with respect to fractional Brownian motion ([56], [52]) and more general Gaussian processes ([47]). Moreover, fbm belongs to an important class of stochastic processes, namely, the Hermite processes. This class appears in non-central limit theorems for processes defined as integrals or partial sums of non-linear functionals of stationary Gaussian sequences with long-range dependence (see [57]). They admit the following representation for all $d \geq 1$:

$$\forall t > 0 \quad Y_t^{H,d} = c(H_0) \int_{\mathbb{R}} ... \int_{\mathbb{R}} \left(\int_0^t \prod_{j=1}^d (s - x_j)_+^{H_0-1} ds \right) dB_{x_1}...dB_{x_d}$$

where $c(H_0)$ is a normalizing constant such that $\mathbb{E}[|Y_1^{H,d}|^2] = 1$ and $H_0 = \frac{1}{2} + \frac{H-1}{d}$ with $H \in (\frac{1}{2}, 1)$. For $d = 1$, one recovers fractional Brownian motion. These processes share many properties with fbm. Namely, they are H-self-similar processes with stationary increments. They possess the same covariance structure, exhibit long range-dependence and their sample paths are almost-surely δ-Hölder continuous, for every $\delta < H$. For $d = 2$, the process is called the Rosenblatt process. This process has received lots of interest in the past and more recent years. Stochastic calculus with respect to the Rosenblatt process has been developed in [73] from both, the pathwise type calculus and Malliavin calculus points of view. Even if these two approaches are successful in order to define a stochastic integral with respect to the Rosenblatt process, the Malliavin calculus one fails to give an Itô's formula for the Rosenblatt process in the divergence sense. In [42], by means of white noise distribution theory, we obtain the following result:

Theorem: Let $(a, b) \in \mathbb{R}_+^*$ such that $a \leq b < \infty$. Let F be an entire analytic function of the complex variable verifying:

$$\exists N \in \mathbb{N}, \exists C > 0, \forall z \in \mathbb{C} \quad |F(z)| \leq C(1 + |z|)^N \exp(\frac{1}{\sqrt{2}b^H}|\Im(z)|)$$

Then, we have in $(S)^*$:

$$F(X_b^H) - F(X_a^H) = \int_a^b F^{(1)}(X_t^H) \diamond \dot{X}_t^H dt + \sum_{k=2}^\infty \left(H\kappa_k(X_1^H) \int_a^b \frac{t^{Hk-1}}{(k-1)!} F^{(k)}(X_t^H) dt + 2^k \int_a^b F^{(k)}(X_t^H) \diamond \dot{X}_t^{H,k} dt \right),$$

where $\{X_t^H\} = \{Y_t^{H,2}\}$, $\{\dot{X}_t^H\}$ is the Rosenblatt noise, $\{\kappa_k(X_1^H); k \geq 2\}$ the non-zero cumulants of the Rosenblatt distribution, $\diamond$ the Wick product and $\{\{X_t^{H,k}\} : k \geq 2\}$ a sequence of processes defined by:

$$\forall t \geq 0 \quad X_t^{H,k} = \int_{\mathbb{R}} \int_{\mathbb{R}} \underbrace{(...((f_t^H \otimes_1 f_t^H) \otimes_1 f_t^H)... \otimes_1 f_t^H)}_{k-1 \times \otimes_1}(x_1, x_2) dB_{x_1} dB_{x_2}$$

with $f_t^H(x_1, x_2) = c(H) \int_0^t \prod_{j=1}^2 (s - x_j)_+^{\frac{H}{2}-1} ds$ and $\otimes_1$ is the contraction of order 1.

Moreover, in the same setting, we obtain the following "isometry" result for the Rosenblatt noise integral of sufficiently "good" integrand processes:

Theorem: Let $\{\phi_t; t \in I\}$ be a stochastic process such that for all $t \in I$ (I an interval), $\phi_t \in (L^2)$ and such that the Rosenblatt noise integral of $\{\phi_t\}$ exists in $(S)^*$. Moreover, let us assume that:

$$\sum_{m=0}^{+\infty} (m+2)! \int_I \int_I |t-s|^{2(H-1)} < f_m(., t); f_m(., s) >_{L^2(\mathbb{R}^m)} dtds < +\infty,$$

where $\phi_t = \sum_{m=0}^{+\infty} I_m(f_m(., t))$. Thus, we have:

$$\mathbb{E}[(\int_I \phi_t \diamond \dot{X}_t^H dt)^2] = H(2H-1) \int_I \int_I |t-s|^{2(H-1)} \mathbb{E}[\phi_t \phi_s] ds dt$$

$$+ 4\sqrt{\frac{H(2H-1)}{2}} \int_I \int_I |t-s|^{H-1} \mathbb{E}[D_{\sqrt{d(H)}\delta_s \circ I_+^{\frac{H}{2}}}(\phi_t) D_{\sqrt{d(H)}\delta_t \circ I_+^{\frac{H}{2}}}(\phi_s)] ds dt$$

$$+ \int_I \int_I \mathbb{E}[(D_{\sqrt{d(H)}\delta_s \circ I_+^{\frac{H}{2}}})^2(\phi_t)(D_{\sqrt{d(H)}\delta_t \circ I_+^{\frac{H}{2}}})^2(\phi_s)] ds dt,$$

where $D_{\sqrt{d(H)}\delta_s \circ I_+^{\frac{H}{2}}}$ is the derivative operator in the direction $\sqrt{d(H)}\delta_s \circ I_+^{\frac{H}{2}}$.

Finally, in the last section of [42], we compare our approach to the one of [73]. More specifically, we prove that the stochastic integral with respect to the Rosenblatt process built using Malliavin calculus corresponds with the Rosenblatt noise integral when both of them exist.

Proposition: Let $\{\phi_t; t \in [0; T]\}$ be a stochastic process such that $\phi \in L^2(\Omega; \mathcal{H}) \cap L^2([0, T]; \mathbb{D}^{2,2})$ and $\mathbb{E}[\int_0^T \int_0^T ||D_{s_1, s_2}\phi||_{\mathcal{H}}^2] ds_1 ds_2 < \infty$ where

$$\mathcal{H} = \{f : [0; T] \to \mathbb{R}; \int_0^T \int_0^T f(s)f(t)|t-s|^{2H-2} ds dt < \infty\}.$$

Then, $\{\phi_t\}$ is Skorohod integrable and $(S)^*$-integrable with respect to the Rosenblatt process, $\{Z_t^H\}_{t \in [0; T]}$, and we have:

$$\int_0^T \phi_t \delta Z_t^H = \int_0^T \phi_t \diamond \dot{Z}_t^H dt$$

6.2. Sample path properties of multifractional Brownian motion

Participants: Paul Balança, Erick Herbin [supervision].

In [50], we have investigated the geometry of the sample paths of multifractional Brownian motion. Several representations of mBm exist, including the classic integral form:

$$\forall t \in \mathbf{R}; \quad X_t = \frac{1}{\Gamma(H(t) + \frac{1}{2})} \int_{\mathbf{R}} \left[(t-u)_+^{H(t)-1/2} - (-u)_+^{H(t)-1/2}\right] dW_u,$$

where $H : \mathbf{R} \mapsto (0, 1)$ is a continuous function. Interestingly, we observe that geometric properties obtained in the probabilistic literature usually rely on a key assumption on the behaviour of the Hurst function:

$$H \text{ is a } \beta\text{-Hölder continuous function such that } \forall t \in \mathbf{R}, \ H(t) < \beta. \qquad (\mathcal{H}_0) \qquad (15)$$

Under the previous hypothesis, the local regularity of the mBm at t corresponds to the geometry of a fractional Brownian motion of parameter $H(t)$. Nevertheless, it has been shown in [15] that when this assumption does not hold, the sample path properties are not as simple and straightforward. More precisely, the latter has proved that the Hölder exponents satisfy at every $t \in \mathbf{R}$:

$$\alpha_{X,t} = H(t) \wedge \alpha_{H,t} \quad \text{and} \quad \widetilde{\alpha}_{X,t} = H(t) \wedge \widetilde{\alpha}_{H,t} \quad \text{a.s.} \qquad (16)$$

This result has been recently improved in [48], observing that the pointwise exponent can even be random under some assumptions on H.

Therefore, the main goal of this work was to obtain a more complete characterization of the geometry of the general mBm. We have first focused on the Hölder regularity of the sample paths, using for this purpose a deterministic representation of the fractional Brownian field:

$$B^{\pm}(t, H) = \frac{\pm 1}{\Gamma\left(H - \frac{1}{2}\right)} \int_{\mathbf{R}} B_u \left[(t - u)_{\pm}^{H-3/2} - (-u)_{\pm}^{H-3/2} \right] \mathrm{d}u, \qquad (17)$$

where $H \geq 1/2$ and B is a continuous Brownian motion. Hence, observing that the mBm almost surely corresponds to the fractional integration of a Brownian motion, we have been able to use the 2-microlocal formalism and its interesting connections with fractional operators. As a consequence, we have proved that the pointwise exponent of the mBm almost surely satisfies:

$$\forall t \in \mathbf{R}; \quad \alpha_{X,t} = H(t) \wedge m_{t,H(t)} \alpha_{H,t}, \qquad (18)$$

where $m_{t,H(t)}$ is defined as the multiplicity of the fractional Brownian field at (t, H), i.e.

$$m_{t,H} = \inf \left\{ k \in \mathbf{N} \smallsetminus \{0\} : \partial_H^k B(t, H) \neq 0 \right\}.$$

We have also been able to obtain some uniform lower bounds on the 2-microlocal frontier, which are optimal under some mild assumptions on the Hurst function.

The second direction of our study has concerned the fractal dimension of the graph of the mBm. Interestingly, and on the contrary to fBm, we have to distinguish the Box and Hausdorff dimensions in our result. The first happens to be the easiest one to study and is closely related to the geometry of H itself. Therefore, with probability one,

$$\forall t \in \mathbf{R} \smallsetminus \{0\}; \quad \dim_{\mathbf{B},t} \mathrm{Gr}(X) = (2 - H(t)) \vee \dim_{\mathbf{B},t} \mathrm{Gr}(H), \qquad (19)$$

where $\dim_{\mathbf{B},t}$ denotes the localized Box dimension at t.

To study the Hausdorff dimension the graph, we need a slightly different approach which makes use of parabolic Hausdorff dimension. We first define for all $t \in \mathbf{R}$ a *parabolic metric* ϱ_H on $\mathbf{R}^2$, with $H > 0$: $\varrho_H\left((u, x); (v, y)\right) := \max\left(|u - v|^H, |x - y|\right)$. For any set $A \subset \mathbf{R}^2$, we denote by $\dim_{\mathcal{H}}\left(A; \varrho_H\right)$ the *parabolic Hausdorff dimension* of A. It is defined similarly to the classic Hausdorff dimension using covering balls relatively to the metric ϱ_H, i.e. it corresponds to the infimum of $s \geq 0$ for which

$$\liminf_{\delta \to 0} \left\{ \sum_{i=0}^{\infty} \operatorname{diam}\left(O_i; \varrho_H\right)^s : (O_i)_{i \in N} \text{ is a } \delta\text{-cover of } A \right\} < \infty$$

Studying the local Hausdorff dimension of the graph of the mBm, we have proved that with probability one

$$\forall t \in \mathbf{R} \smallsetminus \{0\}; \quad \dim_{\mathrm{H},t} \operatorname{Gr}(X) = 1 + H(t)\left(\dim_{\mathrm{H},t}\left(\operatorname{Gr}(H); \varrho_{H(t)}\right) - 1\right). \tag{20}$$

Even though this result might seem counter-intuitive, it can be checked that it induced the classic equality $\dim_{\mathrm{H},t} \operatorname{Gr}(X) = 2 - H(t)$ when the mBm satisfies the assumption $\mathcal{H}_0$. Interestingly, we observe that a similar expression has also emerged recently in the study [70] of the Hausdorff dimension of a fractional Brownian motion with variable drift. Finally, we also note this result can also been extended to images of fractal sets by the multifractional Brownian motion.

6.3. Large Deviations Inequalities

Participant: Xiequan Fan.

Let $(\xi_i)_{i=1,\dots,n}$ be a sequence of independent and centered random variables satisfying Bernstein's condition, for a constant $\varepsilon > 0$,

$$|\mathbb{E}\xi_i^k| \leq \frac{1}{2} k! \varepsilon^{k-2} \mathbb{E}\xi_i^2, \quad \text{for all } k \geq 2 \text{ and all } i = 1, \dots, n. \tag{21}$$

Denote by

$$S_n = \sum_{i=1}^{n} \xi_i \quad \text{and} \quad \sigma^2 = \sum_{i=1}^{n} \mathbb{E}\xi_i^2. \tag{22}$$

The well-known Bernstein inequality (1946) states that, for all $x > 0$,

$$\mathbb{P}(S_n > x\sigma) \leq \inf_{\lambda \geq 0} \mathbb{E}e^{\lambda(S_n - x\sigma)}. \tag{23}$$

In the i.i.d. case, Cramér (1938) has established a large deviation expansion under the condition $\mathbb{E}e^{|\xi_1|} < \infty$. For all $0 \leq x = o\left(\sqrt{n}\right)$, one has

$$\frac{\mathbb{P}(S_n > x\sigma)}{1 - \Phi(x)} = e^{\frac{x^3}{\sqrt{n}}\lambda\left(\frac{x}{\sqrt{n}}\right)}\left[1 + O\left(\frac{1+x}{\sqrt{n}}\right)\right], \quad n \to \infty, \tag{24}$$

where $\lambda(\cdot) = c_1 + c_2 \frac{x}{\sqrt{n}} + \dots$ is the Cramér series and the values $c_1, c_2, \dots$ depend on the distribution of ξ_1.

Bahadur-Rao (1960) proved the following sharp large deviations similar to (15). Assume Cramér's condition. Then, for given $y > 0$, there is a constant c_y depending on the distribution of ξ_1 and y such that

$$\mathbb{P}\left(\frac{S_n}{n} > y\right) = \frac{\inf_{\lambda \geq 0} \mathbb{E}e^{\lambda(S_n - yn)}}{\sigma_y t_y \sqrt{2\pi n}}\left[1 + O\left(\frac{c_y}{n}\right)\right], \quad n \to \infty, \tag{25}$$

where t_y, σ_y and c_y depend on the distribution of ξ_1 and y.

We present an improvement on Bernstein's inequality. In particular, we establish a sharp large deviation expansion similar to the classical results of Cramér and Bahadur-Rao. The following theorem is our main result.

Theorem 0.1 Assume Bernstein's condition. Then, for all $0 \leq x < \frac{1}{12}\frac{\sigma}{\varepsilon}$,

$$\mathbb{P}(S_n > x\sigma) = \inf_{\lambda \geq 0} \mathbb{E}e^{\lambda(S_n - x\sigma)} F\left(x, \frac{\varepsilon}{\sigma}\right), \tag{26}$$

where $\sqrt{2\pi}M(x)$ is the Mills ratio, the function

$$F\left(x, \frac{\varepsilon}{\sigma}\right) = M(x) + 28\,\theta R\left(4x\varepsilon/\sigma\right)\frac{\varepsilon}{\sigma} \tag{27}$$

with

$$R(t) = \frac{\left(1 - t + 6t^2\right)^3}{\left(1 - 3t\right)^{3/2}\left(1 - t\right)^7}, \qquad 0 \leq t < \frac{1}{3}, \tag{28}$$

and $|\theta| \leq 1$. In particular, in the i.i.d. case, for all $0 \leq x = o(\sqrt{n}), n \to \infty$,

$$\left|\mathbb{P}(S_n > x\sigma) - M(x)\inf_{\lambda \geq 0}\mathbb{E}e^{\lambda(S_n - x\sigma)}\right| = O\left(\frac{1}{\sqrt{n}}\inf_{\lambda \geq 0}\mathbb{E}e^{\lambda(S_n - x\sigma)}\right) \tag{29}$$

and thus

$$\frac{\mathbb{P}(S_n > x\sigma)}{M(x)\inf_{\lambda \geq 0}\mathbb{E}e^{\lambda(S_n - x\sigma)}} = 1 + o\left(1\right). \tag{30}$$

6.4. A fractional Brownian field indexed by L^2 and a varying Hurst parameter

Participant: Alexandre Richard.

Using structures of Abstract Wiener Spaces and their reproducing kernel Hilbert spaces, we define a fractional Brownian field indexed by a product space $(0, 1/2] \times L^2(T, m)$, where the first coordinate corresponds to the Hurst parameter of fractional Brownian motion. This field encompasses a large class of existing fractional Brownian processes, such as Lévy fractional Brownian motion and multiparameter fractional Brownian motion, and provides a setup for new ones. We prove that it has good incremental variance in both coordinates and derive certain continuity and Hölder regularity properties. Then, we apply these general results to multiparameter and set-indexed processes, which proves the existence of processes with prescribed local Hölder regularity on general indexing collections.

The family of fBm can be considered for the different Hurst parameters as a single Gaussian process indexed by $(h, t) \in (0, 1) \times \mathbb{R}_+$, which is the position we adopt. Besides, the "time" indexing is replaced by any separable L^2 space. We prove that there exists a Gaussian process indexed by $(0, 1/2] \times L^2(T, m)$, with the additional constraint that the variance of its increments is as well behaved as it was on $(0, 1) \times \mathbb{R}_+$, that is, for any compact of L^2, there is a constant $C > 0$ such that for any f in this compact, and any $h, h' \in (0, 1/2)$,

$$\mathbb{E}\left(B_f^h - B_f^{h'}\right)^2 \leq C\,(h - h')^2. \tag{31}$$

When looking at the L^2–fBf with a fixed h, we have the following covariance: for each $h \in (0, 1/2]$,

$$k_h : (f, g) \in L^2 \times L^2 \mapsto \frac{1}{2} \left(m(f^2)^{2h} + m(g^2)^{2h} - m(|f - g|^2)^{2h} \right) . \tag{32}$$

An important subclass of these processes is formed by processes restricted to indicator functions of subsets of T. In particular, multiparameter when $(T, m) = (\mathbb{R}^d_+, \text{Leb.})$, and more largely set-indexed processes [62],[20] naturally appear and thus motivate generalization *b)*, besides the inherent interest of studying processes over an abstract space.

To define this field, we used fractional operators on the Wiener space W introduced in [56], and first expressed the fractional Brownian field (indexed by $(0, 1/2] \times \mathbb{R}_+$) as a white noise integral over W:

$$\left\{ \int_W \langle \mathcal{K}_h R_h(\cdot, t), w \rangle \, d\mathbb{B}_w, \ (h, t) \in (0, 1/2] \times \mathbb{R}_+ \right\} ,$$

The advantage of this approach is to allow the transfer of techniques of calculus on the Wiener space to any other linearly isometric space with the same structure (those spaces are called Abstract Wiener Spaces). Using the separability and reproducing kernel property of the Cameron-Martin spaces built from the kernels $k_h, h \in (0, 1/2]$, we prove the existence of a Brownian field $\{\mathbf{B}_{h,f}, \ h \in (0, 1/2], f \in L^2(T, m)\}$ over some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Some Hilbert space analysis then provides the desired bound (22). Then, we used this to derive a sufficient condition for almost sure continuity of the fractional Brownian field, in terms of metric entropy.

For fixed h, we proved that the h-fractional Brownian motion has the strong local nondeterminism property, which allowed to compute a sharp estimate of its small deviations, that is, for a compact K of L^2:

$$\exp\left(-C \ N(K, d_h, \varepsilon)\right) \le \mathbb{P} \left(\sup_{f \in K} |\mathbf{B}^h_f| \le \varepsilon \right) \le \exp\left(-C^{-1} \ N(K, d_h, \varepsilon)\right) ,$$

where $N(K, d_h, \varepsilon)$ is the metric entropy of K, i.e., the minimal number of balls necessary to cover K with d_h-balls (the metric induced by the h-fBm) of radius at most ε.

Finally, we looked at the Hölder regularity of the fBf, when the L^2 indexing collection is restricted to the indicator functions of the rectangles of $\mathbb{R}^d$ (multiparameter processes) or to some indexing collection (in the sense of [62]). This restriction permits to use local Hölder regularity exponents, in the flavour of what was done in [24]. When a regular path $\mathbf{h} : L^2 \to (0, 1/2]$ is specified, this defines a multifractional Brownian field as $\mathbf{B}^\mathbf{h}_f = \mathbf{B}_{\mathbf{h}(f),f}$, whose Hölder regularity at each point is proved to equal $\mathbf{h}(f)$ almost surely.

6.5. Self-stabilizing processes

Participants: Xiequan Fan, Jacques Lévy Véhel.

In collaboration with K. Falconer, University of St Andrews.

Self-stabilizing processes are càdlàg processes whose local intensiy of jumpd depend on amplitude. We have investigated two paths to define such processes. The first one is based on a modification of the celebrated Lévy construction of Brownian motion.

The second one starts from a stochastic differentiel equation, and allows one to build Markov processes, a useful feature in applications such as financial modelling [40], [41].

6.6. Multifractal spectra of multistable Lévy motion

Participant: Jacques Lévy Véhel.

In collaboration with R. Le Guével, University of Rennes.

As a follow-up to the work in [34] we have computed the Hausdorf, large deviation, and Legendre multifractal spectra of multistable Lévy motion. It turns out that the shape of the Hausdorf multifractal spectrum is much more complex than could be expected considering the corresponding spectrum of plain Lévy motion. Also, the large deviation spectrum reveals more information on the fine structure of the process than the Hausdorf one, a situation reminiscent of what has already been observed for the model we have developped previously for TCP traffic [2],[39].

6.7. Self-regulating processes for the modelling of geophysical signals

Participant: Jacques Lévy Véhel.

In collaboration with A. Echelard and A. Philippe, University of Nantes.

We have shown that various geophysical signals, and in particular temperature records, can be modelled with self-regulating processes as introduced in [4]. For this purpose, we have used an estimator of the self-regulating function proposed in [44]. Such a modelling allows one to gain further insight on the fine structure of the evolution of temperatures.

6.8. Regularity-preserving signal denoising

Participant: Jacques Lévy Véhel.

In collaboration with A. Echelard.

We have proposed a new wavelet-based method for signal denoising, that allows one to recover the local Hölder regularity of the original signal under weak assumptions [43]. The algotihm is a modification of the well-known wavelet thresholding procedure, where "small" coefficients are not put to zero, but modified in a way governed by the behaviour of large scale coefficients. This will have applications in the frame of our Tandem project on the analysis of radar images.

<div align="center">

SECSI Project-Team

</div>

6. New Results

6.1. Dishonest keys (Objective 2)

Participants: Hubert Comon-Lundh, Guillaume Scerri.

One of the main issues in the formal verification of the security protocols is the validity (and scope) of the formal model. Otherwise, it may happen that a protocol is proved and later someone finds an attack. This paradoxical situation may happen when the formal model used in the proof is too abstract.

A main stream of research therefore consists in proving full abstraction results (also called *soundness*): if the protocol is secure in the (symbolic) model, then an attack can only occur with negligible probability in a computational model. Such results have two main drawbacks: first they are very complicated, and have to be completed again and again for each combination of security primitives. Second, they require strong hypotheses on the primitives, some of which are not realistic. For instance, it is assumed that the attacker cannot forge his own keys (or that all keys come with their certificates, even for symmetric encryption keys).

Hubert Comon-Lundh, Véronique Cortier and Guillaume Scerri had proposed an extension of the symbolic model in 2012, and proved it computationally sound, without this restriction on the dishonest keys.

6.2. Deciding trace equivalence

Participants: David Baelde, Stéphanie Delaune, Rémy Chrétien, Lucca Hirschi.

Most existing results focus on trace properties like secrecy or authentication. There are however several security properties, which cannot be defined (or cannot be naturally defined) as trace properties and require the notion of indistinguishably. Typical examples are anonymity, privacy related properties or statements closer to security properties used in cryptography.

In the framework of the applied pi-calculus as in similar languages based on equational logics, indistinguishably corresponds to a relation called trace equivalence. Roughly, two processes are trace equivalent when an observer cannot see any difference between the two processes. Static equivalence applies only to observations on finite sets of messages, and do not take into account the dynamic behavior of a process whereas the notion of trace equivalence is more general and takes into account this aspect.

6.2.1. Static equivalence.

As explained above, static equivalence is a cornerstone to provide decision procedures for observational equivalence.

Stéphanie Delaune, in collaboration with Mathieu Baudet and Véronique Cortier, has designed a generic procedure for deducibility and static equivalence that takes as input any convergent rewrite system [12]. They have shown that their algorithm covers most of the existing decision procedures for convergent theories. They also provide an efficient implementation. This paper is a journal version of the work presented at RTA'09.

6.2.2. Trace equivalence.

When the processes under study do not contain replication, trace equivalence can be reduced to the problem of deciding symbolic equivalence [13]. Thanks to this reduction and relying on a result first proved by M. Baudet, this yields the first decidability result of observational equivalence for a general class of equational theories (for processes without else branches and without replication). Moreover, based on another decidability result for deciding equivalence between sets of constraint systems, we get decidability of trace equivalence for processes with else branch for standard primitives.

Even though there are some implementations of the procedures described above, this does not suffice to obtain practical tools. Current prototypes suffer from a classical combinatorial explosion problem caused by the exploration of many interleavings in the behaviour of processes. David Baelde, Stéphanie Delaune, and Lucca Hirschi revisit a work due to Mödersheim et al., generalize it and adapt it for equivalence checking. They obtain an optimization in the form of a reduced symbolic semantics that eliminates redundant interleavings on the fly. This work will be published as:

- D. Baelde, S. Delaune, and L. Hirschi. A Reduced Semantics for Deciding Trace Equivalence using Constraint Systems. In *Proc. 3rd Conference on Principles of Security and Trust (POST 2014)*, Grenoble, April 2014, France.

When processes under study contain replication, the approach relying on symbolic equivalence does not work anymore. Moreover, since it is well-known that deciding reachability properties is undecidable under various restrictions, there is actually no hope to do better for equivalence-based properties. Rémy Chrétien, Véronique Cortier, and Stéphanie Delaune provide the first results of (un)decidability for certain classes of protocols for the equivalence problem. They consider a class of protocols shown to be decidable for reachability properties, and establish a first undecidability result. Then, they restrained the class of protocols a step further by making the protocols deterministic in some sense and preventing it from disclosing secret keys. This tighter class of protocols was then shown to be decidable after reduction to an equivalence between deterministic pushdown automata. This work has been published at ICALP'13 [14].

To deal with replication, another approach has been studied by Vincent Cheval in collaboration with Bruno Blanchet. They propose an extension of the automatic protocol verifier ProVerif. ProVerif can prove observational equivalence between processes that have the same structure but differ by the messages they contain. In order to extend the class of equivalences that ProVerif handles, they extend the language of terms by defining more functions (destructors) by rewrite rules. These extensions have been implemented in ProVerif and allow one to automatically prove anonymity in the private authentication protocol by Abadi and Fournet. This work is part of Vincent Cheval's PhD thesis, and was published as:

- V. Cheval, B. Blanchet. Proving More Observational Equivalences with ProVerif. In *2nd Conference on Principles of Security and Trust (POST 2013)*. David Basin, John Mitchell, eds. Springer Verlag, Lecture Notes in Computer Science 7796, 2013.

6.3. Mobile ad-hoc networks

Participants: Rémy Chrétien, Stéphanie Delaune.

Mobile ad hoc networks consist of mobile wireless devices which autonomously organize their communication infrastructure: each node provides the function of a router and relays packets on paths to other nodes. Finding these paths in an a priori unknown and constantly changing network topology is a crucial functionality of any ad hoc network. Specific protocols, called *routing protocols*, are designed to ensure this functionality known as *route discovery*. Secured versions of routing protocols have been proposed to provide more guarantees on the resulting routes, and some of them have been designed to protect the privacy of the users.

Rémy Chrétien and Stéphanie Delaune propose a framework for analysing privacy-type properties for routing protocols. They use the notion of equivalence between traces to formalise three security properties related to privacy, namely indistinguishability, unlinkability, and anonymity. They study the relationship between these definitions and we illustrate them using two versions of the ANODR routing protocol. This work was published as:

- R. Chrétien, S. Delaune. Formal Analysis of Privacy for Routing Protocols in Mobile Ad Hoc Networks. *Principles of Security and Trust - Second International Conference, POST 2013*, held as Part of the *European Joint Conferences on Theory and Practice of Software, ETAPS 2013*, Rome, Italy, March 16-24, 2013. Proceedings. Springer 2013. Lecture Notes in Computer Science. ISBN 978-3-642-36829-5. Pages 1-20.

6.4. Composition results

Participant: Stéphanie Delaune.

Formal methods have proved their usefulness for analysing the security of protocols. However, protocols are often analysed in isolation, and this is well-known to be not sufficient as soon as the protocols share some keys.

Stéphanie Delaune, in collaboration with Céline Chevalier, Steve Kremer, and Mark Ryan, study whether password protocols can be safely composed, even when a same password is reused. More precisely, they present a transformation which maps a password protocol that is secure for a single protocol session (a decidable problem) to a protocol that is secure for an unbounded number of sessions. Their result provides an effective strategy to design secure password protocols: (i) design a protocol intended to be secure for one protocol session; (ii) apply their transformation and obtain a protocol which is secure for an unbounded number of sessions. Their technique also applies to compose different password protocols allowing one to obtain both inter-protocol and inter-session composition. This work was published as:

- C. Chevalier, S. Delaune, S. Kremer and M. Ryan. Composition of Password-based Protocols. *Formal Methods in System Design* 43(3), pages 369-413, 2013.

6.5. Unconditional Soundness (Objective 2)

Participants: Hubert Comon-Lundh, Guillaume Scerri.

Hubert Comon-Lundh, Véronique Cortier and Guillaume Scerri had shown in a 2012 CCS paper how one could drop one of the assumptions of computational soundness results. However, the proofs remain very complicated and there are still assumptions such as the absence of key cycles, or no dynamic corruption... that are still necessary for all these results.

Gergei Bana and Hubert Comon-Lundh investigated a completely different approach to formal security proofs in a 2012 POST paper, which does not make any such assumptions. The idea can be stated in a nutshell: whereas all existing formal models state the attacker's abilities, they propose to formally state what the attacker *cannot* do.

This makes a big difference, since the soundness need only to be proved formula by formula and only the very necessary assumptions are used for such formulas (for instance, no absence of key cycles is needed). This does not need to be proved again when a primitive is added.

Once the general setting is fixed, the question was how practical is the method. We studied the complexity of the consistency proofs in this setting and showed that we can complete such proofs in Polynomial Time for a wide class of axioms in

- H. Comon-Lundh, V. Cortier and G. Scerri. Tractable inference systems: an extension with a deducibility predicate. In CADE'13, LNAI 7898, pages 91-108. Springer, 2013

The development of a prototype implementation is under development. We expect to complete experiments on a number of protocols.

6.6. Static Analysis of Programs with Imprecise Probabilities

Participant: Jean Goubault-Larrecq [correspondant].

Static analyses allows one to obtain guarantees about the behavior of programs, without running them. Programs that handle numerical data such as feedback control loops pose a challenge in this area. This gets even harder when one considers programs that read numerical data from sensors, and write to actuators, as these data are imprecise, and are governed by probability distributions that may themselves be unknown, and only know to fall into some interval of distributions.

As part of the ANR projet blanc CPP, an efficient static analysis framework that deals with this kind of programs was proposed in 2011 by J. Goubault-Larrecq, O. Bouissou, E. Goubault, Sylvie Putot, based on P-boxes and Dempster-Shafer structures to handle imprecise probabilities.

The semantic foundations were made clearer, a new, improved algorithm was proposed, and new applications were examined in:

- A. Adjé, O. Bouissou, J. Goubault-Larrecq, E. Goubault and S. Putot. Static Analysis of Programs with Imprecise Probabilistic Inputs. In *VSTTE'13*, LNCS. Springer, 2013.

<p align="center" style="color:red">SELECT Project-Team</p>

6. New Results

6.1. Model selection in Regression and Classification

Participants: Gilles Celeux, Serge Cohen, Jairo Cugliari, Tim Van Erwen, Clément Levrard, Erwan Le Pennec, Pascal Massart, Nelo Molter Magalhaes, Lucie Montuelle, Mohammed Sedki.

Erwan Le Pennec is still working with Serge Cohen (IPANEMA Soleil) on hyperspectral image segmentation based on a spatialized Gaussian Mixture Model. Their scheme is supported by some theoretical investigation and have been applied in pratice with an efficient minimization algorithm combining EM algorithm, dynamic programming and model selection implemented with MIXMOD. Lucie Montuelle is studying extensions of this model that comprise parametric logistic weights and regression mixtures.

Unsupervised segmentation is an issue similar to unsupervised classification with an added spatial aspect. Functional data is acquired on points in a spatial domain and the goal is to segment the domain in homogeneous domain. The range of applications includes hyperspectral images in conservation sciences, fMRi data and all spatialized functional data. Erwan Le Pennec and Lucie Montuelle are focusing on the questions of the way to handle the spatial component from both the theoretical and the practical point of views. They study in particular the choice of the number of clusters. Furthermore, as functional data require heavy computation, they are required to propose numerically efficient algorithms. They have also extend the model to regression mixture.

Lucie Montuelle focused on conditional density estimation by Gaussian mixtures with logistic weights. Using maximum likelihood estimators, a model selection procedure has been applied, supported by a theoretical guarantee. Numerical experiments have been conducted for regression mixtures with parametric logistic weights, using EM and Newton algorithms. This work is available in the research report and a submitted article.

In collaboration with Lucien Birgé (Université Paris 6), Pascal Massart and Nelo Molter Magalhaes define for the algorithm selection problem a new general cross validation procedure based on robust tests, which is an extension of the hold-out defined by Birgé. They get an original procedure based on the Hellinger distance. This procedure is the unique procedure which does not use any contrast function since it does not estimate the risk. They provide theoretical results showing that, under some weak assumptions on the considered statistical methods, the selected estimator satisfies an oracle type inequality. And, they prove that their robust method can be implemented with a sub-quadratic complexity. Simulations show that their estimator performs generally well for estimating a density with different sample sizes and can handle well-known problems, such as histogram or bandwidth selection.

In collaboration with Gérard Biau (Université Paris 6), Clément Levrard and Pascal Massart provide intuitive conditions have been derived for the k-means clustering algorithm to achieve its optimal rate of convergence. They can be thought of as margin conditions such as ones introduced by Mammen and Tsybakov in the statistical learning framework. These conditions can be checked in many cases, such as Gaussian mixtures with a known number of components and do not require the underlying distribution to have a density, on the contrary to the previous fast rates conditions introduced in this domain. Moreover, It allows to derive non-asymptotic bounds on the mean squared distortion of the k-mean estimator, emphasizing the role played by several other parameters of the quantization issue, such as the smallest distance between optimal codepoints or the excess risk of local minimizers. The influence of these parameters is still in discussion, but some previous results show that some of them are crucial for the minimax results obtained in quantization theory.

Tim van Erven is studying model selection for the long term. When a model selection procedure forms an integrated part of a company's day-to-day activities, its performance should be measured not on a single day, but on average over a longer period, like for example a year. Taking this long-term perspective, it is possible to aggregate model predictions optimally even when the data probability distribution is so irregular that no statistical guarantees can be given for any individual day separately. He studies the relation between model selection for individual days and for the long term, and how the geometry of the models affects both. This work has potential applications in model aggregation for the forecasting of electrical load consumption at EDF. Together with Jairo Cugliari it has also been applied to improve regional forecasts of electrical load consumption using the fact that the consumption of all regions together must add up to the total consumption over the whole country.

The well-documented and consistent variable selection procedure in model-based cluster analysis and classification, that Cathy Maugis (INSA Toulouse) has 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 and Mohammed Sedki, in collaboration with Cathy Maugis, proposed to sort the variables using a lasso-like penalization adapted to the Gaussian mixture model context. Using this rank to select the variables they avoid the combinatory problem of stepwise procedures. Their algorithm is now tested on several challenging simulated and real data sets, showing encouraging performances.

In collaboration with Jean-Michel Marin (Université de Montpellier) and Olivier Gascuel (LIRMM), Gilles Celeux has started a research aiming to select a short list of models rather a single model. This short list of models 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. And, the Kullback-Leibler distances at hand are estimated trough parametric bootstrap procedures.

6.2. Statistical learning methodology and theory

Participants: Vincent Brault, Gilles Celeux, Christine Keribin, Erwan Le Pennec, Lucie Montuelle, Mesrob Ohannessian, Michel Prenat, Solenne Thivin.

Gilles Celeux, Christine Keribin and the Ph D. student Vincent Brault continued their study on the Latent Block Model (LBM), and worked more especially on categorical data. They further investigated a Gibbs algorithm to avoid solutions with empty clusters on synthetic as well as real data (Congressional Voting Records and genomic data) [STCO13].They detailed the link between the information criteria ICL and BIC, compared them on synthetic and real data, and conjectured that these criteria are both consistent for LBM, which is not a standard behavior. ICL has been proved to be preferred for LBM.

V. Brault applied the Large Gaps algorithm and compared it with other existing algorithms [Aussois13]. He also derived a CEM algorithm for categorical LBM [Agroselect13]. In partnership with the Inria- MODAL team, he implemented the algorithms and information criteria in the R package blockcluster.

C. Keribin has started a collaboration with Tristan Mary-Huard (AgroParisTech) by the supervision of an internship (Master 2) on the use of LBM with truncated Poisson data.

Erwan Le Pennec is supervising Solenne Thivin in her CIFRE with Michel Prenat and Thales Optronique. The aim is target detection on complex background such as clouds or sea. Their approach is a local approach based on test decision theory. They have obtained theoretical and numerical results on a segmentation based approach in which a simple Markov field testing procedure is used in each cell of a data driven partition.

Erwan Le Pennec and Michel Prenat have also collaborated on a cloud texture modeling using a non-parametric approach. Such a modeling coud be used to better calibrate the detection procedure: it can lead to more examples than the one acquired and it could be the basis of an ensemble method.

Mesrob Ohannessian joined SELECT through an ERCIM Alain Bensoussan fellowship. During his stay, his work focused on two different aspects of statistics: large datasets and data scarcity. In collaboration with researchers in ETH Zurich (Prof. Andreas Krause), he studied the possibility of trading off statistical performance and computational speed in the context of k-means clustering, using the notion of coresets. In

collaboration with researchers in Paris 11 (Prof. Elisabeth Gassiat) and Paris 7 (Prof. Stéphane Boucheron), he worked on adaptive universal compression when the alphabet is very large, meaning that some symbol observations are scarce.

6.3. Reliability

Participants: Yves Auffray, Gilles Celeux, Rémy Fouchereau, Patrick Pamphile.

Since 2011, in the framework of a CIFRE convention with Snecma-SAFRAN Rémy Fouchereau has started a thesis on the modeling of fatigue lifetime supervised by Gilles Celeux and Patrick Pamphile. In aircraft, space and nuclear industry, fatigue test is the main basic tool for analyzing fatigue lifetime of a given material, component, or structure. A sample of the material is subjected to cyclic loading S (stress, force, strain, etc.), by a testing machine which counts N, the number of cycles to failure. Fatigue test results are plotted on a SN-curve. A probabilistic model for the construction of SN-curve is proposed. In general, fatigue test results are widely scattered for High Cycle Fatigue region and "duplex" SN-curves appears for Very High Cycle region. That is why classic models from mechanic of rupture theory on one hand, probability theory on the other hand, do not fit SN-curve on the whole range of cycles. We have proposed a probabilistic model, based on a fracture mechanic approach: few parameters are required and they are easily interpreted by mechanic or material engineers.This model has been applied to both simulated and real fatigue test data sets. The SN-curves have been well fitted on the whole range of cycles. The parameters have been estimated using the EM algorithm, combining Newton-Raphson optimisation method and Monte Carlo integral estimations. Recently, the model has been improved taking into account production process information, thanks to a clustering approach. Thus, we have provided engineers with a probabilistic tool for reliability design of mechanical parts, but also with a diagnostic tool for material elaboration.

Since 2013, Gilles Celeux and Patrick Pamphile supervise, in the framework of a collaboration with CEA not yet finalized, a thesis on the modeling of battery State Of Charge for electrical vehicles. Electrical battery is an electrochemical device that converts stored chemical energy into electrical energy. This conversion is reversible and can be repeated during charge/discharge cycles. In an electric vehicle, the battery State Of Charge (SOC) gives the driver indication of how long he can drive without recharging the battery. Unfortunately the complex nature of electrochemical reactions does not allow to measure the SOC directly. Different methods of estimation exist, but they are not robust to various environment conditions (temperature, vehicle driving,...) and to the battery ageing. We propose to estimate the SOC from an *Markov-switching model* : the measurement equation specifies how the SOC depends of an unobservable Markov chain and physical data (temperature, voltage and current intensity,...). Moreover, the SOC estimation is included in the Battery Management System, and therefore estimations must be done online, i.e. with minimum information.

A collaboration has started in 2013 with Dassault Aviation on modal analysis of mechanical structures, which aims at identifying the vibration behavior of structures under dynamic excitations. From algorithmic view point, modal analysis amounts to estimation in parametric models on the basis of measured excitations and structural responses data. As it appears from literature and existing implementations, the model selection problem attached to this estimation is currently treated by a rather heavy and very heuristic procedure. The model selection via penalization tools are intended to be tested on this model selection problem.

6.4. Statistical analysis of genomic data

Participants: Vincent Brault, Gilles Celeux, Christine Keribin.

In collaboration with Florence Jaffrezic and Andrea Rau (INRA, animal genetic department), Mélina Gallopin has started a thesis under the supervision of Gilles Celeux. This thesis is concerned with building statistical networks of genes in animal grenetic. In animal genetic, datasets have a large number of genes and low number of statistical units. For this reason, standard network inference techniques work poorly in this case. At first, this team has developed a data-based method to filter replicated RNA-seq experiments. The method, implemented in the Bioconductor R package `HTSFilter`, removes low expressed genes by optimizing the Jaccard index and reduce the dimension of the dataset. Now, they are studying a clustering model on their expression

profiles measured by RNAseq data using Poisson mixture models. External biological knowledge, such as Gene Ontology annotations are taken into account in the model selection step, based on a approximation of the completed log-likelihood given the annotations.

In collaboration with Marie-Laure Martin-Magniette (URGV), GIlles Celeux and Christine Keribin has started a research concerning the buliding statistical networks of transcription factors (TF) with Gaussian Graphical Models (GGM) in the frawork of the intership of Yann Vasseur (Université Paris-sud) who is starting a PhD. thesis on the same subject at the end of 2013. Since the number of TF is greater than the number of statistical units, a lasso-like procedure is used. Moreover the edges of the network are interpreted using the Latent Block Model studied by Vincent Brault in his thesis. An open issue to be solved is the choice of the regularization parameter in the lasso procedure. It is also important to develop this statistical inference for data with good biological control and knowledge to assess the biological relevance of the proposed models.

6.5. Curves classification, denoising and forecasting

Participants: Jairo Cugliari, Émilie Devijver, Pascal Massart, Jean-Michel Poggi, Vincent Thouvenot.

In collaboration with Farouk Mhamdi and Meriem Jaidane (ENIT, Tunis, Tunisia), Jean-Michel Poggi proposed a method for trend extraction from seasonal time series through the Empirical Mode Decomposition (EMD). Experimental comparison of trend extraction based on EMD, X11, X12 and Hodrick Prescott filter are conducted. First results show the eligibility of the blind EMD trend extraction method. Tunisian real peak load is also used to illustrate the extraction of the intrinsic trend.

Jean-Michel Poggi was the supervisor (with A. Antoniadis) of the PhD Thesis of Jairo Cugliari-Duhalde which takes place in a CIFRE convention with EDF. It was strongly related to the use of wavelets together with curves clustering in order to perform accurate load consumption forecasting. The thesis contains methodological and applied aspects linked to the electrical context as well as theoretical ones by introducing external variables in the context of nonparametric forecasting time series. See http://hal.archives-ouvertes.fr/docs/00/78/82/49/PDF/cugliari-jma.pdf and http://hal.inria.fr/docs/00/55/99/39/PDF/RR-7515.pdf The industrial post-doc of Jairo Cugliari, funded by EDF, explores three aspects of this model that complement the original methodology: first, the construction of a confidence interval for the predictor function, second, the flexibility and simplicity of the model to provide, without extra effort, forecasts horizons further and further away and finally, and third: study of the ability to provide good predictions in the presence of subtle signal nonstationarities induced by loss of customers coming from various scenarios, see http://hal.archives-ouvertes.fr/docs/00/81/49/24/PDF/kwf-suite.pdf

Jean-Michel Poggi, co-supervising with Anestis Antoniadis (Université Joseph Fourier Grenoble) the PhD thesis of Vincent Thouvenot, funded by a CIFRE with EDF. The industrial motivation of this work is the recent development of new technologies for measuring power consumption by EDF to acquire consumption data for different mesh network. The thesis will focus on the development of new statistical methods for predicting power consumption by exploiting the different levels of aggregation of network data collection. From the mathematical point of view, the work is to develop generalized additive models for this type of kind of aggregated data for the modeling of functional data, associating closely nonparametric estimation and variable selection using various penalization methods.

Jean-Michel Poggi and Pascal Massart are the co-advisors of the PhD thesis of Émilie Devijver, strongly motivated by the same kind of industrial forecasting problems in electricity, is dedicated to curves clustering for the prediction. A natural framework to explore this question is mixture of regression models for functional data. The theoretical subject of the thesis is to extend to functional data the recent work by Bühlmann et al. dealing with the simultaneous estimation of mixture regression models in the scalar case using Lasso type methods. Of course, it will be based on the technical tools of the work of Caroline Meynet (which completes her thesis Orsay under the direction of P. Massart), which deals with the clustering of functional data using Lasso methods choosing simultaneously number of clusters and selecting significant wavelet coefficients.

6.6. Neuroimaging, Statistical analysis of fMRI data

Participants: Gilles Celeux, Christine Keribin.

This research takes place as part of a collaboration with Neurospin on brain functional Magnetic Resonance Imaging (fMRI) data. (http://www.math.u-psud.fr/select/reunions/neurospin/Welcome.html). and concerns essentially regularisation in a supervised clustering methodology that includes spatial information in the prediction framework, and yields clustered weighted maps. C. Keribin examined the PhD defence of Virgile Fritsch High-dimensional statistical methods for inter-subjects studies in neuroimaging (Inria, Parietal team).

Specfun Team

6. New Results

6.1. Creative telescoping for bivariate hyperexponential functions

In [8], we gave a new algorithm for the symbolic integration of bivariate hyperexponential functions, which outperforms state-of-the-art implementations like Maple's function *DEtools[Zeilberger]*. The approach was to extend Hermite's reduction for rational functions and the Hermite-like reduction for hyperexponential functions in a suitable way. A key feature of the algorithm is that it can avoid the costly computation of certificates.

6.2. Creative telescoping for rational functions

In [10] we described a precise and elementary algorithmic version of the Griffiths–Dwork method for the creative telescoping of rational functions. This leads to bounds on the order and degree of the coefficients of the differential equation, and to the first complexity result which is single exponential in the number of variables. One of the important features of the algorithm is that it does not need to compute certificates. The approach is vindicated by a prototype implementation.

6.3. Complexity of the uncoupling of linear functional systems

Uncoupling algorithms transform a linear differential system of first order into one or several scalar differential equations. We examined in [9] two approaches to uncoupling: the cyclic-vector method (*CVM*) and the Danilevski-Barkatou-Zürcher algorithm (*DBZ*). We gave tight size bounds on the scalar equations produced by *CVM*, and designed a fast variant of *CVM* whose complexity is quasi-optimal with respect to the output size. We exhibited a strong structural link between *CVM* and *DBZ* enabling to show that, in the generic case, *DBZ* has polynomial complexity and that it produces a single equation, strongly related to the output of *CVM*. We proved that algorithm *CVM* is faster than *DBZ* by almost two orders of magnitude, and provided experimental results that validate the theoretical complexity analyses.

6.4. Computation of integrals related to the Ising model

We showed in [2] that the n-fold integrals of the magnetic susceptibility of the Ising model, as well as various other n-fold integrals of the "Ising class", or n-fold integrals from enumerative combinatorics, like lattice Green functions, correspond to a distinguished class of functions generalising algebraic functions: they are actually diagonals of rational functions. This algebraic structure explains many remarkable properties of the integrals of the Ising class.

6.5. Non-D-finite excursions in the quarter plane

The number of excursions (finite paths starting and ending at the origin) having a given number of steps and obeying various geometric constraints is a classical topic of combinatorics and probability theory. We proved in [3] that the sequence of numbers of excursions in the quarter plane corresponding to a nonsingular step set $S \subseteq \{0, \pm 1\}^2$ with infinite group does not satisfy any nontrivial linear recurrence with polynomial coefficients. Accordingly, in those cases, the trivariate generating function of the numbers of walks with given length and prescribed ending point is not D-finite. This solves an open problem in the field of lattice path combinatorics.

6.6. A human proof of Gessel's lattice path conjecture

Gessel walks are planar walks confined to the positive quarter plane, that move by unit steps in any of the following directions: West, North-East, East, and South-West. In 2001, Ira Gessel conjectured a closed-form expression for the number of Gessel walks of a given length starting and ending at the origin. In 2008, Kauers, Koutschan, and Zeilberger gave a computer-aided proof of this conjecture. The same year, Bostan and Kauers showed, using again computer algebra tools, that the trivariate generating function of Gessel walks is algebraic. We proposed in [15] the first "human proofs" of these results. They are derived from a new expression for the generating function of Gessel walks.

6.7. Efficient algorithms for rational first integrals

We presented in [14] fast algorithms for computing rational first integrals with bounded degree of a planar polynomial vector field. Our approach is inspired by an idea of Ferragut and Giacomini. We improve upon their work by proving that rational first integrals can be computed via systems of linear equations instead of systems of quadratic equations. This leads to a probabilistic algorithm with arithmetic complexity $\tilde{O}(N^{2\omega})$ and to a deterministic algorithm solving the problem in $\tilde{O}(d^2 N^{2\omega+1})$ arithmetic operations, where N denotes the given bound for the degree of the rational first integral, and where $d \leq N$ is the degree of the vector field, and ω the exponent of linear algebra. By comparison, the best previous algorithm uses at least $d^{\omega+1} N^{4\omega+4}$ arithmetic operations. The new algorithms are very efficient in practice.

6.8. Reactive document checking in Coq

In an effort to improve the reactivity of Coq, the way it processes and checks a single document has been completely redesigned [7]. The current development version is able to reschedule the tasks to be performed in order to minimize the time required to give interactive feedback to the user. On typical documents taken from the formal proof of the Odd Order Theorem, the worst reaction time of the tool dropped from 5 minutes to 9 seconds. This improvement will be part of the next stable release of the Coq system.

6.9. Efficient normalization of ring/field expressions in Coq

The implementation of Coq's proof commands for manipulation of ring/field expressions has been improved in response to the demand for better efficiency that emerged in the formalization of Apéry's irrationality proof of $\zeta(3)$. The data structure used for the abstract syntax tree of ring/field expressions has been refined to enable a more efficient and more precise interpretation into concrete ring/field expressions. Moreover the collection of non-nullity conditions for denominators in a field expressions has been speeded up, making the type-checking time of a field normalization proof not be dominated by this collecting phase.

6.10. Documentation of Coq's canonical structures

The device employed to model a hierarchy of algebraic structures with overloaded notations in Coq has been documented in [6] and in the user manual of the tool.

6.11. Maintenance and development of the SSReflect extension for Coq and its user manual

The Small Scale Reflection extension of Coq has been maintained together with its user manual. Some new linguistic constructs to model non-structural reasoning and to enable the user to better factor out repeated arguments have been developed and documented. Some language constructs have been made compatible with the type-classes mechanism offered by Coq. The release of version 1.5 has been prepared.

6.12. Efficient proof-search techniques in sequent calculus

We have proposed in [11] a sequent calculus which is focussed, polarized, and parameterized by an abstract notion of theory. This new combination of features aims at proposing a framework which is adapted to the simulation in sequent calculus of efficient, general-purpose decision procedures (tableaux methods, satisfiability, ...) that can interact with theory-specific decision procedures (for linear arithmetics, arrays, ...). In particular we propose a tight simulation of the Davis–Putnam–Logemann–Loveland algorithm modulo theory, and show how to simulate some advanced optimizations that are crucial to realistic implementations of SMT solvers.

6.13. A formal proof of the irrationality of $\zeta(3)$

We have obtained a formal proof, machine-checked by the Coq proof assistant, of the irrationality of the constant $\zeta(3)$, under the single assumption of the asymptotic behavior of the least common multiple of the first n natural numbers. The core of this formal proof is based on (untrusted) computer-algebra calculations performed outside the proof assistant with the Algolib Maple library. Then, we verify formally and a posteriori the desired properties of the objects computed by Maple and complete the proof of irrationality.

6.14. Documentation of the Mathematical Components libraries

The approach to finite-group theory adopted in the libraries formalizing in Coq the proof of the Odd Order Theorem has been documented in [5].

TAO Project-Team

6. New Results

6.1. Continuous Optimization

Participants: Ouassim Ait Elhara, Yohei Akimoto, Asma Atamna, Anne Auger, Alexandre Chotard, Nikolaus Hansen, Ilya Loshchilov, Yann Ollivier, Marc Schoenauer, Michèle Sebag, Olivier Teytaud.

Our expertise in continuous optimization is focused on stochastic search algorithms. We address theory, algorithm design, and applications. The methods we investigate are adaptive techniques that are able to learn iteratively the parameters of the distribution used to sample (new) solutions. The Covariance Matrix Adaptation Evolution Strategy (CMA-ES) is nowadays one of the most powerful methods for derivative-free continuous optimization. We work on different variants of the CMA-ES to improve it in various contexts as described below. We have previously proven the convergence of simplified variants of the CMA-ES algorithm using the theory of stochastic approximation, and have provided the first proofs of convergence on composite of twice continuously differentiable functions. More recently, we used Markov chain analysis for analyzing the step-size adaptation rules of evolution strategies related to the CMA-ES algorithm.

Surrogate models for CMA-ES. In the context of his PhD thesis defended in January 2013 [4], Ilya Loshchilov has proposed different surrogate variants of CMA-ES based on ranking-SVM that preserve the invariance to monotonic transformation of the CMA-ES algorithm. As a follow-up, he has proposed an original over-exploitation mechanism in case of accurate surrogate [44]. Several of these models have entered the BBOB-2013 workshop [43]. Further research direction using the KL divergence between successive distributions as a trigger for a new learning phase has been proposed [45].

Step-size adaptive methods. We have proposed a new step-size adaptation mechanism that can loosely be interpreted as a new variant of the 1/5 success rule for comma (non-elitist) strategies and which is applicable with a large population size [21]. The rule uses the success of the median fitness of the current population compared to a (different) fitness percentile from the previous population.

Principles of Stochastic Optimization. Based on the framework of *information geometry* (IGO), theoretical guarantees have been obtained for continuous optimization algorithms: using the *natural gradient* provides improvement guarantees even when using finite step sizes [22]. We have considered the principles of designing effective stochastic optimization algorithms from the bottom-up and the top-down perspective [56]. The top-down perspective takes the information-geometrical view-point and largely confirms the bottom-up construction.

Benchmarking. We have continued our effort for improving standards in benchmarking and pursued the development of the COCO (COmparing Continuous Optimizers) platform (see Section 5.4). We have organized the ACM GECCO 2013 workshop on Black-Box-Optimization Benchmarking [2] and benchmarked different surrogate-based variants of the CMA-ES algorithm [26], [44], [43]. Our new starting ANR project NumBBO, centered on the COCO platform, aims at extending it for large-scale, expensive, constrained and multi-objective optimization.

Theoretical proofs of convergence. We have established the connection between convergence of comparison based step-size adaptive randomized search and the stability analysis of some underlying Markov chains. This connection heavily exploits invariance properties of the algorithm. In a first paper we establish this connection for scaling-invariant functions and prove sufficient conditions for linear convergence expressed in terms of stability conditions [63]. We have proven, using this defined methodology, the linear convergence of a famous algorithm introduced independently by several researchers and known as the (1+1)-ES with one-fifth success rule [62]. In [32], we have proven the linear convergence of a modified evolutionary algorithm without assuming quasi-convexity.

---

[2] see http://coco.gforge.inria.fr/doku.php?id=bbob-2013

6.2. Optimal Decision Making under Uncertainty

Participants: Olivier Teytaud [correspondent], Jean-Joseph Christophe, Adrien Couëtoux, Jérémie Decock, Nicolas Galichet, Manuel Loth, Marc Schoenauer, Michèle Sebag.

The UCT-SIG works on sequential optimization problems, where a decision has to be made at each time step along a finite time horizon, and the underlying problem involves uncertainties along an either adversarial or stochastic setting.

The most prominent application domain is now energy management, at various time scales, and more generally, planning in uncertain environments. The main advances done this year include:

- A work on metagaming/investment [12], where a macroscopic decision has to be made (e.g., investment decisions, which plants should be built) prior to operational decisions (e.g., unit commitment policy, i.e., the operational management of the system). This is a key part of our activity for 2014.

- Bandit problems with risk [36]. Bandit problems are quite related to metagaming problems (they correspond to the unstructured case).

- A theoretical work on the consistency of Monte Carlo Tree Search / Upper Confidence Tree in continuous domains [27]. A non-trivial extension was necessary for proving such a consistency.

- Noisy optimization is a key part of our work [61], as it is crucial for direct policy search or more generally for dynamic optimization:

    – We have proven lower bounds under "locality assumptions" [33], which are usually informally assumed by some practicionners for justifying the use of evolutionary algorithms.

    – In cases with strong noise (variance not decreasing to zero around the optimum) we proved log-log convergence for simple rules for choosing the number of resamplings [23].

- Several submissions are joint works with Ailab, National Dong Hwa University, Hualien, Taiwan. The drafts can be found at http://www.lri.fr/~teytaud/indema.html.

- In collaboration with Christian Shulte (KTH, Stockholm), one of the main contributors to the well-known general-purpose CP solver *GECODE* (http://www.gecode.org/), and within the Microsoft-Inria joint lab Adapt project, ideas from UCT have been integrated in GECODE for the choice of the variable values during the exploration of the constraint tree. The most critical issue lied in the definition of a meaningful reward for a given node (variable = value) that could cope with the multiple restarts of the search: the deeper the failure, the larger the reward (and hence this work also pertains to the CRI-SIG(Section 6.4). Initial results have been obtained with job-shop scheduling problems [47] and more extensive results have been obtained on 3 benchmarks of the CP community [46].

6.3. Distributed Systems

Participants: Cécile Germain-Renaud [correspondent], Philippe Caillou, Dawei Feng, Cyril Furtlehner, Victorin Martin, Michèle Sebag.

The DIS-SIG explores the issues related to modeling and optimizing distributed systems, ranging from very large scale computational grids to multi-agent systems and large scale traffic management.

Fault management. As Lamport formulated decades ago, fault management in distributed systems exemplifies the unreachability of exact prior knowledge. Real-world large scale systems additionally face the non-stationarity issue.

[20] models the system state and its ruptures (non-stationarity) through the flow of jobs as a stream (scalability), with a traceability goal (interpretability), and addresses a key difficulty in Data Streaming, which is timely detection of a change in the generative process underlying the data stream *drift*. A statistical model based on spatial distance and time frequency is proposed, together with adaptive thresholding. Theoretical and experimental validation show the robustness of the method.

D. Feng's PhD formulates the problem of probe selection for fault prediction based on end-to-end probing as a Collaborative Prediction (CP) problem, based on the reasonable assumption of an underlying factorial model [13]. Monitoring large scale distributed systems differ from CP's usual applications (personalized recommendation), in two major ways. On the brighter side, while users cannot be queried for specific recommendations, probes can be launched at will. On the downside, firstly the distribution of the probe results is highly skewed, faults being a small fraction of the total population, and secondly, some of the faults are transient. Amongst the numerous approaches addressing CP, Minimum Margin Matrix Factorization (MMMF) is easily amenable to active learning, which addresses fault sparsity both at spatial (skewed distributions) and temporal (transients) level. From extensive experiments on real-world data, we have shown that modelling probe-based fault prediction as a CP task and addressing this task through MMMF is an extremely efficient strategy for fault prediction. Comparative analysis and experiments motivate the critical advantage of active learning. It offers a scalable alternative to direct AUC optimization. Similarly, comparison with bias-aware methods (Mixed Membership Matrix Factorization) indicates that the capacity of actively selecting the most informative probes provides the most efficient method to capture the time variability of the system.

Multi-agent and games.   Resuming earlier work, our goal is to provide an automated abstract description of simulation results. Data mining methods are used to identify groups in complex simulations [11]. Using activity indicators to identify the most interesting agent groups [17], the groups and their evolution are described through one or several simulations [10]. To facilitate the dissemination of the algorithms, we have participated in the development of a generic multiagent platform (GAMA), in collaboration with IRD, University of Rouen (IDEES), and University of Toulouse (IRIT) [34], [35].

A statistical physics perspective.   With motivating applications in large scale traffic congestion inference problems, we have

- Settled a method for encoding real data with pairwise dependencies into an Ising copula [58] suitable to infer real-valued data (travel time) from the computation of its corresponding latent binary state (congested/not congested) probabilities.

- In parallel we have investigated the inverse Ising problem, proposing among other methods a loop analysis based on a duality transformation, leading to a dual belief propagation algorithm runing on the dual graph formed by the network of independent cycles. This aims at finding an MRF to represent pairwise correlated data, close to a dependence tree, able to take into account most important loops [14], [15].

6.4. Designing Criteria

Participants: Jamal Atif, Yoann Isaac, Mostepha Redouane Khouadjia, Hélène Paugam-Moisy, Marc Schoenauer, Michèle Sebag.

This recently created SIG, rooted on the claim that *What matters is the criterion*, aims at defining new learning or optimization objectives reflecting fundamental properties of the model, the problem or the expert prior knowledge.

Image understanding.   We continued our effort on the development of model-based image understanding approaches (e.g. [71]). In [18], we have proposed a method for simultaneously segmenting and recognizing objects in images, based on a structural representation of the scene and on a constraint propagation method. Theoretical guaranties have been provided along with a quantitative assessment on healthy and pathological brain structures in magnetic resonance images. Within the ANR project LOGIMA (collaboration with ECP, Telecom PariTech and TU Dresden), our goal is to introduce a new lattice-based representation and reasoning framework suited for dealing with spatial objects in the presence of uncertainty. This framework associates under the aegis of general lattice theory ingredients from mathematical morphology, description logics and formal concept analysis. A first development of this framework can be found in [7] where it has been exploited for the definition of

abductive reasoning services and applied to high-level image understanding. Several theoretical issues have been raised in the development of this new framework. Some of them were tackled in [25], [24], [30]. In [25], we have shown how mathematical morphology operators can ben defined on general concept lattices. Explicit join-commuting and meet-commuting operators are defined either from particular valuations on the corresponding lattice or from the decomposition of their elements. In [24], we extended our mathematical morphology based adductive reasoning to multivalued logics, hence allowing us to deal with several uncertainty and imprecision phenomena. In [30], metrics between bipolar information - where the information is represented by a positive/preference part and a negative/constraint part - have been introduced based on particular dilations.

Structured learning.   With motivations in bio-informatics and brain computer interfaces, the goal is to take into account priors about the spatio-temporal structure of the underlying phenomenon in order to propose a generative model of the data.

In the context of Yoann Isaac's PhD (Digiteo Unsupervised Brain project), in collaboration with CEA LIST, the goal is to design a representation endowed with appropriate invariance properties. Specifically, within the framework of sparse dictionary coding, we have introduced new priors allowing us to capture both spatial and temporal regularity of multivariate brainwave signals [54]. The learning/optimization criterion, while being multivariate, contains several non-differentiable terms, raising new optimization issues; the proposed approach extends the classical split Bregman iterations algorithm to the multi-dimensional case with several non-differentiable terms [37].

In the context of regulatory gene networks, the challenge is to combine probabilistic inference (does a gene regulate another one) with relational learning (the set of genes is organized in a network). Ensemble learning approaches have been used to cope with the imbalanced nature of the data, e.g., bagging Markov logic networks or boosting operator-valued kernel-based regressors [55], [64]. Another issue, regarding the indeterminacy of the models due to the data sparsity, is addressed through prior-guided regularization beyond model sparsity such as orthogonality [8] or stability [16].

In the domain of medical imaging, the exploitation of computational tomography data to model tumor physiology is hindered by the huge noise level; the multi-task setting is leveraged to provide a better robustness to noise [51].

Robotic value systems.   Within the European SYMBRION IP, investigations on the preference-based reinforcement learning were continued in Riad Akrour's PhD, where the robot demonstrations are assessed by the expert and these assessments are used to learn a model of the expert's expectations. In [67], this work had been extended and combined with active learning to yield state-of-the art performances with few binary feedbacks from the expert. The work has first concentrated this year on handling the noise due to expert's mistakes [53], and bridging the reality gap when porting the algorithms on real robots (e-pucks and one Nao robot) – these results will be published in Riad Akrour's PhD dissertation, to be defended in Spring 2014.

Algorithm Selection as Collaborative Filtering.   The crucial issue when addressing algorithm selection problems is to be able to come up with features that can describe the problems: with representative features, algorithm selection amounts to supervised learning. However, except for some rare domaines (e.g., SAT, [73]), no satisfying set of features exists. However, algorithm selection can also be viewed as a recommendation problem, and tackled by collaborative filtering: users more or less like movies, and similarly, instances like algorithms as much as these algorithms are efficient in solving it. Applying collaborative filtering leads to designing a latent feature space in which the representation of the problems is highly adapted to the algorithm selection problem. A critical issue in collaborative filtering is the 'cold start' problem, that is making recommendations for a brand new user/problem instance. This issue has been handled by a surrogate model of the latent factors, mapping the initial features onto the latent ones. The *Algorithm Recommender System* has been successfully applied to 3 different domains: Satisfiability, Constraint Programming, and Continuous Black-Box Optimization (data from the COCO platform, see Section 5.4) [59].

Social Networks with insider information.   The analysis of social networks based on the contents and

structure of information exchanges most often pertains to descriptive learning, e.g., explaining the growth of the communication graph or investigating the sensitivity of existing algorithms to hyper-parameters [31]. In the Modyrum context (coll. SME Augure), a supervised learning perspective is investigated, taking advantage of the fact that experts already know part of the sought results in some specific domains of interest. Based on e.g., Twitter and blogs data, the goal is to define generic features and supervised learning algorithms, enabling to characterize the targets of interest depending on the public relation focus.

Multi-objective AI Planning.   Within the ANR project DESCARWIN (http://descarwin.lri.fr), Mostepha-Redouane Kouadjia continued his work on the **multi-objective approach** to AI Planning using the Evolutionary Planner *Divide-and-Evolve* (DaE), that evolves a sequential decomposition of the problem at hand: each sub-problem is then solved in turn by some embedded classical planner [70]. Even though the embedded planner is single-objective, DaE can nevertheless handle multi-objective problems. Current work includes the implementation of the multi-objective version of DaE, the definition of some benchmark suite, and some first numerical experiments, comparing in particular the results of a full Pareto approach to those of the classical aggregation method. These works resulted in 3 conference papers recently accepted, introducing a tunable benchmark test suite [39], demonstrating that the best quality measure for parameter tuning in this multi-objective framework is the hypervolume, even in the case of the aggregation approach [41], and comparing the evolutionary multi-objective approach with the aggregation method, the only method known to the AI Planning community [38]. A sum-up of these recent results have been published at IJCAI [40].

6. New Results

6.1. Deductive Verification

- F. Bobot, J.-C. Filliâtre, C. Marché, G. Melquiond, and A. Paskevich have presented the proof session mechanism of *Why3* at VSTTE 2013 [23]. It is a technique to maintain a proof session against modification of verification conditions. It was successfully used in developing more than a hundred verified programs and in keeping them up to date along the evolution of *Why3* and its standard library. It also helps out with changes in the environment, *e.g.* prover upgrades.

- M. Clochard, C. Marché, and A. Paskevich 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. This work will be presented at the PLPV conference in January 2014 [29]

- M. Clochard published at the POPL conference a paper presenting a work done during an internship at Rice University (Houston, TX, USA) with S. Chaudhuri and A. Solar-Lezama [28]. It is a new technique for parameter synthesis under boolean and quantitative objectives. The input to the technique is a "sketch" — a program with missing numerical parameters — and a probabilistic assumption about the program's inputs. The goal is to automatically synthesize values for the parameters such that the resulting program satisfies: (1) a boolean specification, which states that the program must meet certain assertions, and (2) a quantitative specification, which assigns a real valued rating to every program and which the synthesizer is expected to optimize.

- J.-C. Filliâtre, L. Gondelman, and A. Paskevich have formalized the notion of ghost code implemented in *Why3*, in a paper *The Spirit of Ghost Code* [49] to be submitted. This is an outcome of L. Gondelman's M2 internship (spring/summer 2013).

- In 2013, two public releases of *Why3* were launched, version 0.81 in March and version 0.82 in December [42]. A first important evolution relies on significant efficiency improvements both in terms of execution speed and of memory usage. The second major evolution is the support for many new provers, including interactive provers PVS 6 (used at NASA) and Isabelle2013-2 (planned to be used in the context of Ada program via Spark), and automated ones: CVC4, Mathematica, Metitarski, Metis, Beagle, Princess, and Yices2. The design of the programming language of *Why3* (WhyML) was presented during a tool demonstration at the ESOP conference [33].

6.2. Floating-Point and Numerical Programs

- S. Boldo, F. Clément, J.-C. Filliâtre, M. Mayero, G. Melquiond, and P. Weis, finished the formal proof of a numerical analysis program: the second order centered finite difference scheme for the one-dimensional acoustic wave [15].

- S. Boldo developed a formal proof of an algorithm for computing the area of a triangle, an improvement of its error bound and new investigations in case of underflow [25].

- S. Boldo, J.-H. Jourdan, X. Leroy, and G. Melquiond, extended CompCert to get the first formally verified compiler that provably preserves the semantics of floating-point programs [26].

- S. Boldo and G. Melquiond wrote a chapter of the book [38] that describes the current state of the Mathematics/Computer science research in France.

- C. Lelay worked on formalizing power series for the Coq proof assistant [35].

- Most 18-year old French students pass an exam called Baccalaureate which ends the high school and is required for attending the university. The idea was to try our Coq library Coquelicot on the 2013 mathematics test of the scientific Baccalaureate. C. Lelay went to the "Parc de Vilgénis" high school in Massy, France and took the 2013 test at the same time as the students, but had to formally prove the answers. There was therefore no possible cheating: the Coq library was already developed and it was tested as is during the four hours of the test. This experiment shows that Coquelicot is able to cope with basic real analysis: it has the necessary definitions and lemmas, and its usability and efficiency have been demonstrated in a test with a limited time [45] (see also https://www.lri.fr/~lelay/).

- D. Ishii and G. Melquiond applied methods of deductive program verification to ensure the safety of hybrid automata [34].

- É. Martin-Dorel, G. Hanrot, M. Mayero, L. Théry, showed how to generate and formally check certificates in the Coq proof assistant to solve myriads of instances of the Integer Small Value Problem (ISValP). This problem is directly related to solving the Table Maker's Dilemma with hardest-to-round computations [50]. A new version of the formalized library has been released (http://tamadi.gforge.inria.fr/CoqHensel/).

- É. Martin-Dorel, G. Melquiond, and J.-M. Muller, studied issues related to double rounding in the implementation of error-free transformations [16].

6.3. Automated Reasoning

- C. Dross, S. Conchon, J. Kanig, and A. Paskevich have proposed a new approach for handling quantified formulas in SMT solvers. Their framework is based on the notion of instantiation patterns, also known as triggers, that suggest instances which are more likely to be useful in proof search. This framework has been implemented in the Alt-Ergo SMT solver [48].

- S. Conchon, A. Goel, S. Krstic, A. Mebsout, and F. Zaïdi have designed a new model checking algorithm that is able to infer invariants strong enough to prove complex parameterized cache-coherence protocols [30].

- S. Conchon, A. Mebsout, and F. Zaïdi have presented a new SMT library called Alt-Ergo-Zero. This library is tightly integrated to the backward reachability algorithm of the Cubicle model checker [31].

- S. Conchon, M. Iguernelala, and A. Mebsout have designed a collaborative framework for reasoning modulo simple properties of non-linear arithmetic. This framework has been implemented in the Alt-Ergo SMT solver [47].

- J. C. Blanchette and A. Paskevich 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). This extension, named TFF1, was incorporated in the TPTP standard and was presented at the CADE-24 conference [22].

6.4. Certification of Languages, Tools and Systems

- A. Tafat and C. Marché developed a certified VC generator using Why3. 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 nor 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 very significantly higher amount of proof automation compared to *Coq* [36]

- A. Charguéraud, together with the other members of the *JsCert* team have developed this year the first complete formalization of the semantics of the JavaScript programming language. This project is joint work with Philippa Gardner, Sergio Maffeis, Gareth Smith, Daniele Filaretti and Daiva Naudziuniene from Imperial College, and Alan Schmitt and Martin Bodin from Inria Rennes – Bretagne Atlantique (see http://jscert.org).

The formalization consists of a set of inductive rules translating the prose from the *ECMAScript Language Specification, version 5*. These rules can be used to formally reason about program behaviors or to establish the correctness of program transformations. In addition to the inductive rules, a reference interpreter has been proved correct. This interpreter may be used to run actual JavaScript program following the rules of the formal semantics. It has been used in particular to validate the formal semantics against official JavaScript test suites.

The formalization of JavaScript has been published at POPL 2014 [24]. A key ingredient in this formalization is the use of the *pretty-big-step semantics*. This technique allows for representing evaluation rules in big-step style without suffering from a duplication of several premises across different rules. The pretty-big-step technique is described in a paper published by A. Charguéraud at ESOP 2013 [27].

- É. Contejean, together with V. Benzaken and their PhD student S. Dumbrava, have proposed a *Coq* formalization of the relational data model which underlies relational database systems [21]. Proposing such a formalization is the first, *essential* step, that will allow to *prove* that existing systems conform to their specifications and to *verify* both production implementations of database systems and database-backed applications. More precisely, they present and formalize the data definition part of the model including integrity constraints, attributes, tuples, relations, schemas and integrity constraints (including the so-called Armstrong's system and the chase). They model two different query language formalisms: relational algebra and conjunctive queries. The former is the basis of the SQL commercial query language and the latter underlies graphical languages, such as Microsoft Access or Query By Example (QBE). They also present logical query optimization and prove the main "database theorems": algebraic equivalences, the homomorphism theorem and conjunctive query minimization.

6.5. Miscellaneous

- R. El Sibaie and J.-C. Filliâtre have developed *Combine*, an OCaml library for combinatorics. It provides two different solutions to the exact matrix cover problem: Knuth's dancing links and ZDDs, a variant of binary decision diagrams [32].