Activity Report Saclay - Île-de-France 2018

Edition: 2019-03-11

List of Inria's Research Teams

Project-Team AVIZ

Analysis and Visualization

RESEARCH CENTER
Saclay - Île-de-France

THEME
Interaction and visualization

Table of contents

Project-Team AVIZ

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

Keywords:

Computer Science and Digital Science:

A1.3. - Distributed Systems
A1.3.3. - Blockchain
A3.1.4. - Uncertain data
A3.1.7. - Open data
A3.1.8. - Big data (production, storage, transfer)
A3.3. - Data and knowledge analysis
A3.3.1. - On-line analytical processing
A3.3.3. - Big data analysis
A3.5.1. - Analysis of large graphs
A5.1. - Human-Computer Interaction
A5.1.2. - Evaluation of interactive systems
A5.1.6. - Tangible interfaces
A5.1.8. - 3D User Interfaces
A5.1.9. - User and perceptual studies
A5.2. - Data visualization
A6.3.3. - Data processing

Other Research Topics and Application Domains:

B1. - Life sciences
B1.1. - Biology
B1.2. - Neuroscience and cognitive science
B9.5.6. - Data science
B9.6. - Humanities
B9.6.1. - Psychology
B9.6.3. - Economy, Finance
B9.6.6. - Archeology, History
B9.6.10. - Digital humanities

1. Team, Visitors, External Collaborators

Research Scientists

Jean-Daniel Fekete [Team leader, Inria, Senior Researcher, HDR]
Tobias Isenberg [Team leader, Inria, Senior Researcher, HDR]
Pierre Dragicevic [Inria, Researcher]
Steve Haroz [Inria, Starting Research Position, from Sep 2018]
Petra Isenberg [Inria, Researcher]
Catherine Plaisant [Fondation Inria, Advanced Research Position, from Sep 2018 until Nov 2018]

External Collaborators

Evelyne Lutton [INRA]

Frédéric Vernier [Univ Paris-Sud, from Oct 2018]
Christoph Kinkeldey [Freie Universität Berlin, Germany, from Apr 2018]

Technical Staff

Christoph Kinkeldey [Inria, until Jan 2018]
Christian Poli [Inria]

PhD Students

Marc Barnabe [INRA]
Sarkis Halladjian [Inria]
Mickael Sereno [Inria, from Oct 2018]
Natkamon Tovanich [Institut de recherche technologique System X, from Dec 2018]
Xiyao Wang [Inria]

Post-Doctoral Fellows

Tanja Blascheck [Inria]
Qing Chen [Ecole polytechnique, from Oct 2018]
Paola Tatiana Llerena Valdivia [Inria, from Jun 2018]

Visiting Scientists

Qing Chen [Hong Kong University of Science and Technology, from Apr 2018 until Jun 2018]
Michael Mcguffin [École de Technologie Supérieure, Montréal, from Oct 2018 until Nov 2018]
Claudio Silva [New York University, from Aug 2018]
Wesley Willett [University of Calgary, Canada, from Nov 2018]

Administrative Assistant

Katia Evrat [Inria]

2. Overall Objectives

2.1. Objectives

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

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

Within this research scope Aviz focuses on five research themes:

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

2.2. Research Themes

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

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

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

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

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

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

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

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

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

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

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

3. Research Program

3.1. Scientific Foundations

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

The use of graphic representation to understand abstract data is a goal Visual Analytics shares with Tukey's Exploratory Data Analysis (EDA) [71], graphic designers such as Bertin [60] and Tufte [70], and HCI researchers in the field of Information Visualization [59].

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

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

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

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

Several theories of perception are related to information visualization such as the "Gestalt" principles, Gibson's theory of visual perception [64] and Triesman's "preattentive processing" theory [69]. We use them extensively but they only have a limited accuracy for predicting the effectiveness of novel visual representations in interactive settings.

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

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

In 2002, we showed that the practical limit of InfoVis was on the order of 1 million items displayed on a screen [62]. Although screen technologies have improved rapidly since then, eventually we will be limited by the physiology of our vision system: about 20 millions receptor cells (rods and cones) on the retina. Another problem will be the limits of human visual attention, as suggested by our 2006 study on change blindness in large and multiple displays [61]. Therefore, visualization alone cannot let us understand very large data sets. Other techniques such as aggregation or sampling must be used to reduce the visual complexity of the data to the scale of human perception.

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

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

Designing novel interaction and visualization techniques to explore huge data sets is an important goal and requires solving hard problems, but how can we assess whether or not our techniques and systems provide real improvements? Without this answer, we cannot know if we are heading in the right direction. This is why we have been actively involved in the design of evaluation methods for information visualization [68], [67], [65], [66], [63]. For more complex systems, other methods are required. For these we want to focus on longitudinal evaluation methods while still trying to improve controlled experiments.

3.2. Innovation

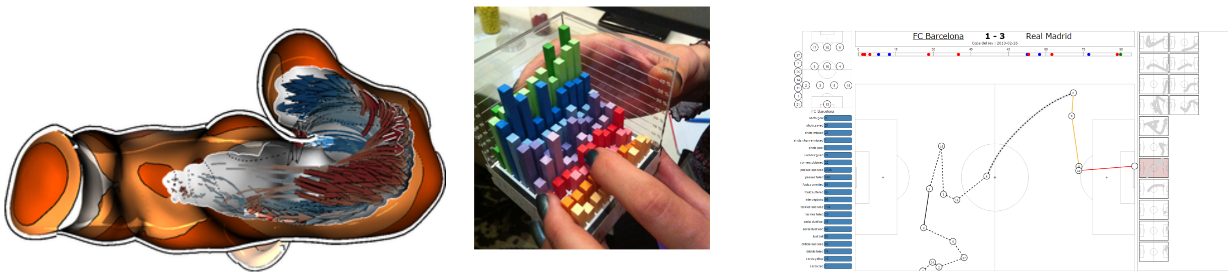


Figure 1. Example novel visualization techniques and tools developed by the team. Left: a non-photorealistic rendering technique that visualizes blood flow and vessel thickness. Middle:a physical visualization showing economic indicators for several countries, right: SoccerStories a tool for visualizing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

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

3.3. Evaluation Methods

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

3.4. Software Infrastructures

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

3.5. Emerging Technologies

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

We want to empower humans to make use of data using different types of display media and to enhance how they can understand and visually and interactively explore information. This includes novel display equipment and accompanying input techniques. The Aviz team specifically focuses on the exploration of the use of large displays in visualization contexts as well as emerging physical and tangible visualizations. In terms of interaction modalities our work focuses on using touch and tangible interaction. Aviz participates to the Digiscope project that funds 11 wall-size displays at multiple places in the Paris area (see http://www.

digiscope.fr), connected by telepresence equipment and a Fablab for creating devices. Aviz is in charge of creating and managing the Fablab, uses it to create physical visualizations, and is also using the local wall-size display (called WILD) to explore visualization on large screens. The team also investigates the perceptual, motor and cognitive implications of using such technologies for visualization.

3.6. Psychology

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

4. Highlights of the Year

4.1. Highlights of the Year

- Steve Haroz joined Aviz as a research scientist (SRP) for three years.
- Catherine Plaisant joined Aviz as an International Chair for 5 years.
- The team welcomed two invited professors (Claudio Silva and Michael McGuffin).
- Aviz members presented seven papers at IEEE VIS 2018 and won a best paper award at Eurovis 2018.
- Former Aviz PhD student Lonni Besançon received a PhD thesis prize honorable mention award from GDR, AFIG, AFRV, and EGFR for his thesis “An interaction Continuum for 3D Data Visualization.”
- Aviz started an Associated Team with the ilab at the University of Calgary on the topic of Situated and Embedded Visualization.

5. New Software and Platforms

5.1. Cartolabe

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

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

5.2. BitConduite

BitConduite Bitcoin explorer
KEYWORDS: Data visualization - Clustering - Financial analysis - Cryptocurrency

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

- Contact: Petra Isenberg

6. New Results

6.1. Declarative Rendering Model for Multiclass Density Maps

Participants: Jaemin Jo [Dept. of Computer Science and Engineering, Seoul National University, South Korea], Pierre Dragicevic, Jean-Daniel Fekete [correspondent].

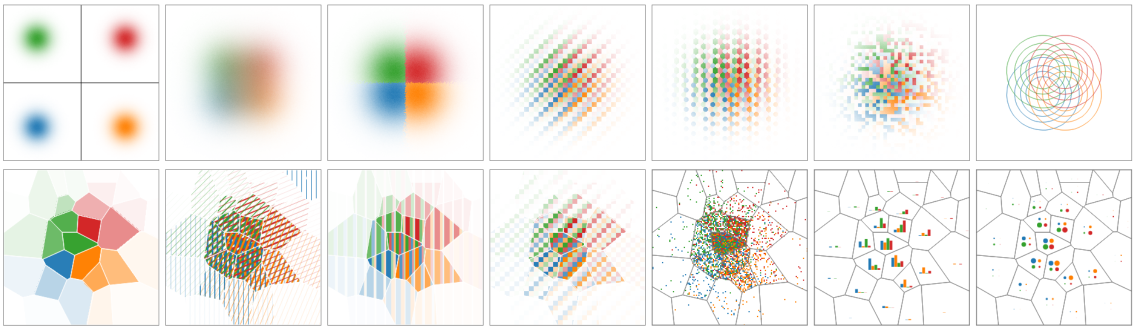


Figure 3. Design alternatives for a four-class density map.

Multiclass maps are scatterplots, multidimensional projections, or thematic geographic maps where data points have a categorical attribute in addition to two quantitative attributes. This categorical attribute is often rendered using shape or color, which does not scale when overplotting occurs. When the number of data points increases, multiclass maps must resort to data aggregation to remain readable. We use a novel model called *multiclass density maps*: multiple 2D histograms computed for each of the category values. Multiclass density maps are meant as a building block to improve the expressiveness and scalability of multiclass map visualization. This library implements our declarative model: a simple yet expressive JSON grammar associated with visual semantics, that specifies a wide design space of visualizations for multiclass density maps. Our declarative model is expressive and can be efficiently implemented in visualization front-ends such as modern web browsers. Furthermore, it can be reconfigured dynamically to support data exploration tasks without recomputing the raw data. Finally, we demonstrate how our model can be used to reproduce examples from the past and support exploring data at scale.

More on the project page: Multiclass Density Maps.

6.2. Reducing Affective Responses to Surgical Images through Color Manipulation and Stylization

Participants: Lonni Besançon [Linköping University Norrköping, Sweden], Amir Semmo [Hasso Plattner Institute, University of Potsdam, Germany], David Biau [Assistance Publique – Hôpitaux de Paris, France], Bruno Frachet [Assistance Publique – Hôpitaux de Paris, France], Virginie Pineau [Institut Curie, France], El Hadi Sariali [Assistance Publique – Hôpitaux de Paris, France], Rabah Taouachi [Institut Curie, France], Tobias Isenberg, Pierre Dragicevic [correspondant].

Figure 4. One of the surgery filters used in our study.

We presented the first empirical study on using color manipulation and stylization to make surgery images more palatable [38]. While aversion to such images is natural, it limits many people's ability to satisfy their curiosity, educate themselves, and make informed decisions. We selected a diverse set of image processing techniques, and tested them both on surgeons and lay people. While many artistic methods were found unusable by surgeons, edge-preserving image smoothing gave good results both in terms of preserving information (as judged by surgeons) and reducing repulsiveness (as judged by lay people). Color manipulation turned out to be not as effective.

This study is an initial investigation but opens up exciting avenues for future research. These include supporting surgery videos, other types of medical images than open surgery (e.g., skin diseases), as well as disturbing imagery outside the medical domain, such as offensive user-generated content that can psychologically impact professionals who monitor it.

All supplemental material is on the OSF page: osf.io/4pfes/.

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

Participants: Evanthia Dimara [ISIR, Sorbonne Université, France], Anastasia Bezerianos [ISIR, Sorbonne Université, France], Pierre Dragicevic [correspondant].

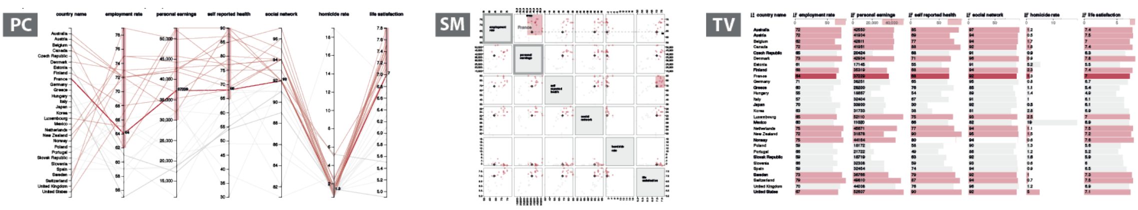


Figure 5. The three visualization techniques tested in our study.

We explored how to rigorously evaluate multidimensional visualizations for their ability to support decision making [22]. We first defined multi-attribute choice tasks, a type of decision task commonly performed with such visualizations. We then identified which of the existing multidimensional visualizations are compatible with such tasks, and evaluated three elementary visualizations: parallel coordinates, scatterplot matrices and tabular visualizations. Our method consisted in first giving participants low-level analytic tasks, in order to ensure that they properly understood the visualizations and their interactions. Participants were then given multi-attribute choice tasks consisting of choosing holiday packages. We assessed decision support through multiple objective and subjective metrics, including a decision accuracy metric based on the consistency between the choice made and self-reported preferences for attributes. We found the three visualizations to be comparable on most metrics, with a slight advantage for tabular visualizations. In particular, tabular visualizations allowed participants to reach decisions faster. Thus, although decision time is typically not central in assessing decision support, it can be used as a tie-breaker when visualizations achieve similar decision accuracy. Our results also suggest that indirect methods for assessing choice confidence may allow to better distinguish between visualizations than direct ones.

All supplemental material is on the project web page: aviz.fr/dm.

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

Participants: Pierre Dragicevic [correspondant], Yvonne Jansen [ISIR, Sorbonne Université, France].

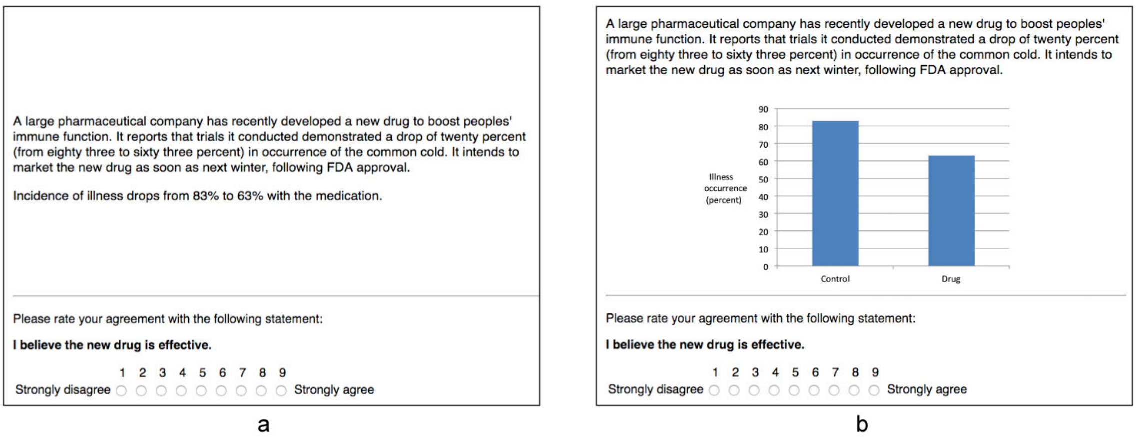


Figure 6. a) text without chart, b) text with "trivial" chart.

We provided a reappraisal of Tal and Wansink's study "Blinded with Science", where seemingly trivial charts were shown to increase belief in drug efficacy, presumably because charts are associated with science. Through a series of four replications conducted on two crowdsourcing platforms, we investigated an alternative explanation, namely, that the charts allowed participants to better assess the drug's efficacy [24]. Considered together, our experiments suggested that the chart seems to have indeed promoted understanding, although the effect is likely very small. Meanwhile, we were unable to replicate the original study's findings, as text with chart appeared to be no more persuasive – and sometimes less persuasive – than text alone. This suggests that the effect may not be as robust as claimed and may need specific conditions to be reproduced. Regardless, within our experimental settings and considering our study as a whole (N = 623), the chart's contribution to understanding was clearly larger than its contribution to persuasion.

The main lesson from our study is that with charts, the peripheral route of persuasion cannot be studied independently from the central route: in order to establish that a chart biases judgment, it is necessary to also rigorously establish that it does not aid comprehension. Our replication also opens many relevant questions for infovis. Are charts really associated with science? More generally, what associations do charts or visualizations trigger depending on their visual design? When exactly is a chart trivial?

All supplemental material is on the project web page: aviz.fr/blinded.

6.5. A Model of Spatial Directness in Interactive Visualization

Participants: Stefan Bruckner [University of Bergen, Norway], Tobias Isenberg [correspondant], Timo Ropinski [Ulm University, Germany], Alexander Wiebel [Hochschule Worms University of Applied Sciences, Germany].

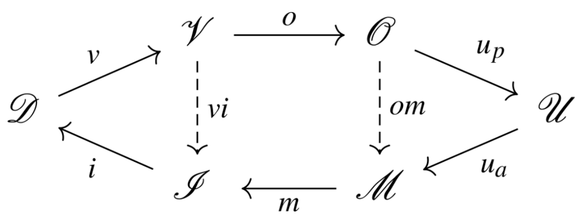


Figure 7. Illustration of the model of spatial directness.

We discussed the concept of directness in the context of spatial interaction with visualization. In particular, we proposeed a model (see Figure 7) that allows practitioners to analyze and describe the spatial directness of interaction techniques, ultimately to be able to better understand interaction issues that may affect usability. To reach these goals, we distinguished between different types of directness. Each type of directness depends on a particular mapping between different spaces, for which we consider the data space, the visualization space, the output space, the user space, the manipulation space, and the interaction space. In addition to the introduction of the model itself, we also showed how to apply it to several real-world interaction scenarios in visualization, and thus discussed the resulting types of spatial directness, without recommending either more direct or more indirect interaction techniques. In particular, we demonstrated descriptive and evaluative usage of the proposed model, and also briefly discussed its generative usage.

More on the project Web page: https://tobias.isenberg.cc/VideosAndDemos/Bruckner2018MSD.

6.6. Multiscale Visualization and Scale-Adaptive Modification of DNA Nanostructures

Participants: Haichao Miao [TU Wien, Austria, and Austrian Institute of Technology, Vienna, Austria], Elisa de Llano [Austrian Institute of Technology, Vienna, Austria], Johannes Sorger [Complexity Science Hub Vienna, Austria], Yasaman Ahmadi [Austrian Institute of Technology, Vienna, Austria], Tadija Kekic [Austrian Institute of Technology, Vienna, Austria], Tobias Isenberg [correspondant], M. Eduard Gröller [TU Wien, Austria], Ivan Barišic [Austrian Institute of Technology, Vienna, Austria], Ivan Viola [TU Wien, Austria and KAUST, Kingdom of Saudi Arabia].

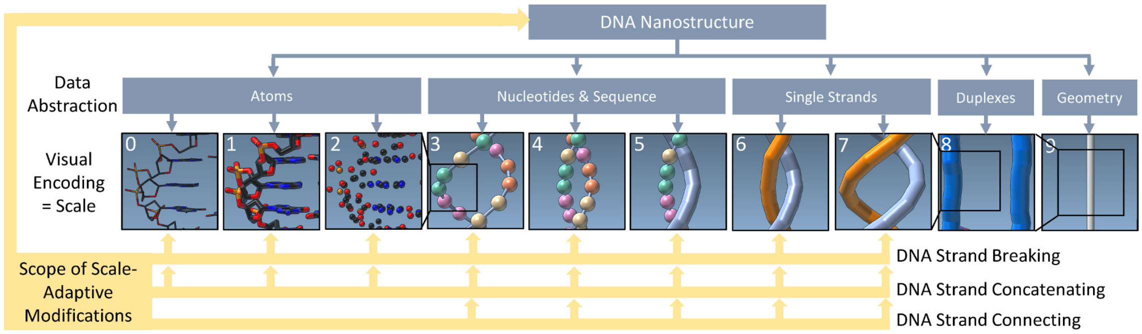


Figure 8. Illustration of the abstraction space.

We presented an approach to represent DNA nanostructures in varying forms of semantic abstraction, describe ways to smoothly transition between them, and thus create a continuous multiscale visualization and interaction space for applications in DNA nanotechnology. This new way of observing, interacting with, and creating DNA nanostructures enables domain experts to approach their work in any of the semantic abstraction levels, supporting both low-level manipulations and high-level visualization and modifications. Our approach allows them to deal with the increasingly complex DNA objects that they are designing, to improve their features, and to add novel functions in a way that no existing single-scale approach offers today. For this purpose we collaborated with DNA nanotechnology experts to design a set of ten semantic scales (see Figure 8). These scales take the DNA's chemical and structural behavior into account and depict it from atoms to the targeted architecture with increasing levels of abstraction. To create coherence between the discrete scales, we seamlessly transition between them in a well-defined manner. We used special encodings to allow experts to estimate the nanoscale object's stability. We also added scale-adaptive interactions that facilitate the intuitive modification of complex structures at multiple scales. We demonstrate the applicability of our approach on an experimental use case. Moreover, feedback from our collaborating domain experts confirmed an increased time efficiency and certainty for analysis and modification tasks on complex DNA structures. Our method thus offers exciting new opportunities with promising applications in medicine and biotechnology.

More on the project Web page: https://tobias.isenberg.cc/VideosAndDemos/Miao2018MVS.

6.7. DimSUM: Dimension and Scale Unifying Maps for Visual Abstraction of DNA Origami Structures

Participants: Haichao Miao [TU Wien, Austria, and Austrian Institute of Technology, Vienna, Austria], Elisa de Llano [Austrian Institute of Technology, Vienna, Austria], Tobias Isenberg [correspondant], M. Eduard Gröller [TU Wien, Austria], Ivan Barišic [Austrian Institute of Technology, Vienna, Austria], Ivan Viola [TU Wien, Austria and KAUST, Kingdom of Saudi Arabia].

We presented a novel visualization concept for DNA origami structures that integrates a multitude of representations into a DimSUM. This novel abstraction map (see Figure 9) provides means to analyze, smoothly transition between, and interact with many visual representations of the DNA origami structures in an effective way that was not possible before. DNA origami structures are nanoscale objects, which are challenging to model in silico. In our holistic approach we seamlessly combined three-dimensional realistic shape models, two-dimensional diagrammatic representations, and ordered alignments in one-dimensional arrangements, with semantic transitions across many scales. To navigate through this large, two-dimensional abstraction map we highlighted locations that users frequently visit for certain tasks and datasets. Particularly interesting viewpoints can be explicitly saved to optimize the workflow. We have developed DimSUM together

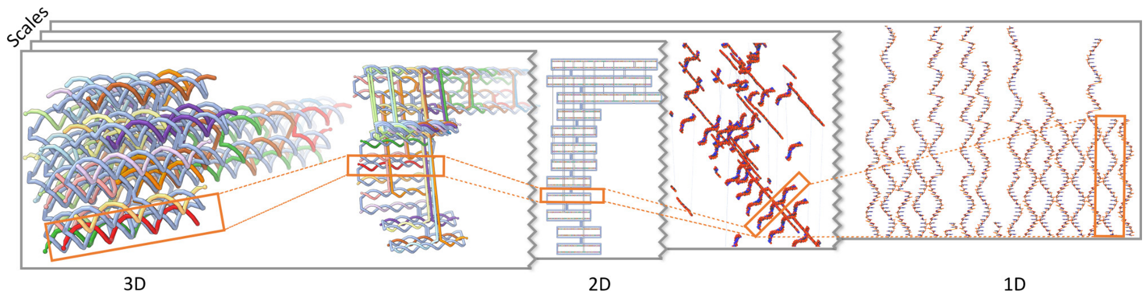


Figure 9. Illustration of the DimSUM space.

with domain scientists specialized in DNA nanotechnology. In the paper we discussed our design decisions for both the visualization and the interaction techniques. We demonstrateed two practical use cases in which our approach increases the specialists' understanding and improves their effectiveness in the analysis. Finally, we discussed the implications of our concept for the use of controlled abstraction in visualization in general.

More on the project Web page: https://tobias.isenberg.cc/VideosAndDemos/Miao2018DDS.

6.8. Pondering the Concept of Abstraction in (Illustrative) Visualization

Participants: Ivan Viola [TU Wien, Austria and KAUST, Kingdom of Saudi Arabia], Tobias Isenberg [correspondant].

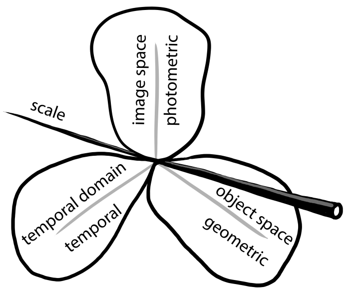


Figure 10. Illustration of the abstraction concept.

We discussed the concept of directness in the context of spatial interaction with visualization (Figure 10). In particular, we proposed a model (autoreffig:directness) that allows practitioners to analyze and describe the spatial directness of interaction techniques, ultimately to be able to better understand interaction issues that may affect usability. To reach these goals, we distinguished between different types of directness. Each type of directness depends on a particular mapping between different spaces, for which we consider the data space, the visualization space, the output space, the user space, the manipulation space, and the interaction space. In addition to the introduction of the model itself, we also showed how to apply it to several real-world interaction scenarios in visualization, and thus discussed the resulting types of spatial directness, without recommending either more direct or more indirect interaction techniques. In particular, we demonstrated descriptive and evaluative usage of the proposed model, and also briefly discussed its generative usage.

More on the project Web page: https://tobias.isenberg.cc/VideosAndDemos/Bruckner2018MSD.

6.9. Is there a reproducibility crisis around here? Maybe not, but we still need to change

Participants: Alex Holcombe [School of Psychology, The University of Sydney], Charles Ludowici [School of Psychology, The University of Sydney], Steve Haroz [correspondant].

Those of us who study large effects may believe ourselves to be unaffected by the reproducibility problems that plague other areas. However, we will argue that initiatives to address the reproducibility crisis, such as preregistration and data sharing, are worth adopting even under optimistic scenarios of high rates of replication success. We searched the text of articles published in the Journal of Vision from January through October of 2018 for URLs (our code is here: https://osf.io/cv6ed/) and examined them for raw data, experiment code, analysis code, and preregistrations. We also reviewed the articles' supplemental material. Of the 165 articles, approximately 12% provide raw data, 4% provide experiment code, and 5% provide analysis code. Only one article contained a preregistration. When feasible, preregistration is important because p-values are not interpretable unless the number of comparisons performed is known, and selective reporting appears to be common across fields. In the absence of preregistration, then, and in the context of the low rates of successful replication found across multiple fields, many claims in vision science are shrouded by uncertain credence. Sharing de-identified data, experiment code, and data analysis code not only increases credibility and ameliorates the negative impact of errors, it also accelerates science. Open practices allow researchers to build on others' work more quickly and with more confidence. Given our results and the broader context of concern by funders, evident in the recent NSF statement that "transparency is a necessary condition when designing scientifically valid research" and "pre-registration... can help ensure the integrity and transparency of the proposed research", there is much to discuss.

6.10. Visualizing Ranges over Time on Mobile Phones: A Task-Based Crowdsourced Evaluation

Participants: Matthew Brehmer [Microsoft Research, USA], Bongshin Lee [Microsoft Research, USA], Petra Isenberg [correspondant], Eun Kyoung Choe [University of Maryland, USA].

In the first crowdsourced visualization experiment conducted exclusively on mobile phones, we experimentally compare approaches to visualizing ranges over time on small displays. People routinely consume such data via a mobile phone, from temperatures in weather forecasting apps to sleep and blood pressure readings in personal health apps. However, we lack guidance on how to effectively visualize ranges on small displays in the context of different value retrieval and comparison tasks, or with respect to different data characteristics such as periodicity, seasonality, or the cardinality of ranges. Central to our experiment is a comparison between two ways to lay out ranges: a more conventional linear layout strikes a balance between quantitative and chronological scale resolution, while a less conventional radial layout emphasizes the cyclicality of time and may prioritize discrimination between values at its periphery. With results from 87 crowd workers, we found that while participants completed tasks more quickly with linear layouts than with radial ones, there were few differences in terms of error rate between layout conditions. We also found that participants performed similarly with both layouts in tasks that involved comparing superimposed observed and average ranges.

More on the project Web page.

7. Partnerships and Cooperations

7.1. Regional Initiatives

- Tobias Isenberg received an equipment grant from STIC, Paris-Saclay, for approx. EUR 5K

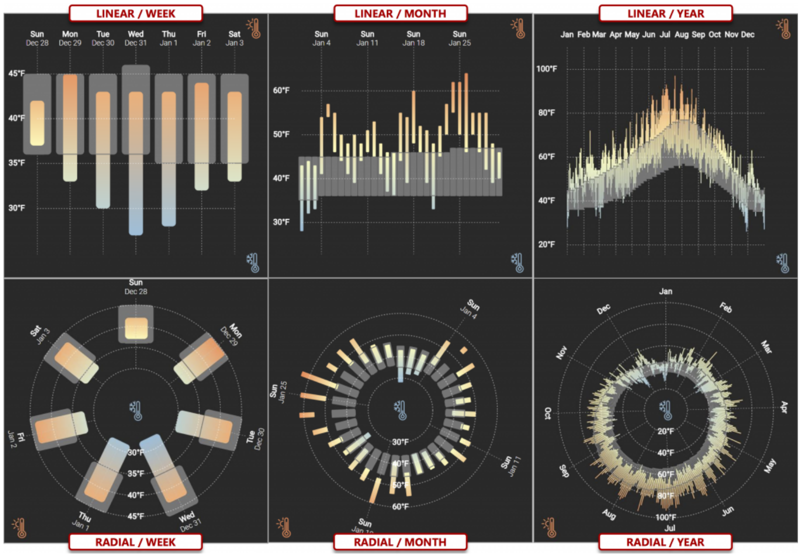


Figure 11. Linear and Radial temperature range charts designed for mobile phone displays, representative of the stimuli used in our crowdsourced experiment. The gradient bars encode observed temperature ranges and are superimposed on gray bars encoding average temperature ranges. Corresponding Week, Month, and Year charts display the same data.

7.2. National Initiatives

- Naviscope Inria Project Lab on Image-guided NAvigation and VIsualization of large data sets in live cell imaging and microSCOPy; collaboration with several Inria project teams and external collaborators; this grant supports a PhD position and funds travel and equipment.

7.3. European Initiatives

7.3.1. FP7 & H2020 Projects

7.3.1.1. IVAN

Title: Interactive and Visual Analysis of Networks

Programm: CHIST-ERA

Duration: May 2018 - April 2021

Coordinator: Dr. Torsten Möller, Uni Wien, Austria

Partners:

EPFL, Switzerland

Inria France

Uni Wien, Austria

Inria contact: Jean-Daniel Fekete

The main goal of IVAN is to create a visual analysis system for the exploration of dynamic or time-dependent networks (from small to large scale). Our contributions will be in three principle areas:

1. novel algorithms for network clustering that are based on graph harmonic analysis and level-of-detail methods;

2. the development of novel similarity measures for networks and network clusters for the purpose of comparing multiple network clusterings and the grouping (clustering) of different network clusterings; and
3. a system for user-driven analysis of network clusterings supported by novel visual encodings and interaction techniques suitable for exploring dynamic networks and their clusterings in the presence of uncertainties due to noise and uncontrolled variations of network properties.

Our aim is to make these novel algorithms accessible to a broad range of users and researchers to enable reliable and informed decisions based on the network analysis.

7.3.2. Collaborations in European Programs, Except FP7 & H2020

- Illustrare project co-funded by ANR, France, and FWF, Austria, funding a PhD position and funds for travel and equipment. The project investigates integrative visual abstraction of molecular data and is a collaboration with TU Wien, Austria

7.4. International Initiatives

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

7.4.1.1. SEVEN

Title: Situated and Embedded Visualization for Data Analysis

International Partner (Institution - Laboratory - Researcher):

University of Calgary (Canada) - ILab - Wesley Willett

Start year: 2018

See also: http://aviz.fr/seven

The goal of this joint work between the Aviz team at Inria Saclay and the ILab at the University of Calgary is to develop and study situated data visualizations to address the limitations of traditional platforms of data analytics. In a situated data visualization, the data is directly visualized next to the physical space, object, or person it refers to. Situated data visualizations can surface information in the physical environment and allow viewers to interpret data in-context, monitor changes over time, make decisions, and act on the physical world in response to the insights gained. However, research on this topic remains scarce and limited in scope. We will build on our track record of successfull collaborations to jointly develop situated visualization as a novel research direction. The objective for the first year is to design and implement situated visualizations to support health and aging. Our joint work is expected generate benefits at multiple levels, including to society and industry (by empowering individuals and professionals with technology), to the scientific community (by developing a new research direction), to the academic partners (by reinforcing existing research links and establishing them as leaders on the topic), and to students (by providing them with unique training opportunities with a diverse team of world-class researchers).

7.4.2. Inria International Partners

7.4.2.1. Informal International Partners

- University of Maryland (USA), University of Roma (Italy), TU Darmstadt (Germany): Jean-Daniel Fekete Fekete collaborates with Leilani Battle, Giuseppe Santucci, Carsten Binnig and colleagues on the design of database benchmarks to better support visualization;
- University of Seoul (Korea): Jean-Daniel Fekete collaborates with Jaemin Jo and Jinwook Seoh on progressive algorithms and visualization techniques;
- University of Bari (Italy): Jean-Daniel Fekete collaborates with Paolo Buono on hypergraph visualization;

- Stanford University. Pierre Dragicevic and Jean-Daniel Fekete collaborate with Sean Follmer on swarm user interfaces.
- Hasso Plattner Institute. Pierre Dragicevic and Tobias Isenberg collaborate with Amir Semmo on stylization filters for facilitating the examination of disturbing visual content.
- University of Minnesota, USA: Tobias Isenberg is collaborating with Daniel F. Keefe on topics of the interactive exploration of 3D data.
- University of Granada, Spain: Tobias Isenberg is collaborating with Domingo Martin and German Arroyo on digital stippling.
- The University of Sydney, Australia. Steve Haroz collaborate with Alex Holcombe on analyzing open practices in vision science.
- Massachusetts Institute of Technology (CSAIL). Steve Haroz collaborates with Aude Oliva on investigating the impact of titles on memory of visualized data.
- University of Washington, University of Zurich and University of Toronto. Pierre Dragicevic and Steve Haroz collaborate with Matthew Kay and Chat Wacharamanotham on transparent statistical reporting and efficient statistical communication. Pierre Dragicevic collaborates with Matthew Kay and Fanny Chevalier on supporting research transparency with interactive research papers.
- University of Calgary. Pierre Dragicevic, Tobias Isenberg, and Petra Isenberg collaborate with Wesley Willett, Sheelagh Carpendale, and Lora Oehlberg on situated data visualization.
- Microsoft Research Redmond and University of Maryland. Petra Isenberg collaborate with Bongshin Lee, Mathieu Brehmer, and Eun Kyoung Choe on Mobile Visualization
- Microsoft Research Redmond. Petra Isenberg and Tanja Blascheck collaborate with Bongshin Lee on Micro Visualizations for Smartwatches

7.5. International Research Visitors

7.5.1. Visits of International Scientists

- Claudio Silva (August 2018 – June 2019): Sabbatical from New York University (USA). Also, invited professor through a DigiCosme grant for 3 months. Claudio Silva is spending one year with Aviz. We launched a bi-weekly seminar on explainable machine-learning with visualization.
- Michael McGuffin (October – November): visit from ETS Montreal (Canada). Michael McGuffin has spent a month with Aviz working on augmented reality and visualization, collaborating with Pierre Dragicevic, Jean-Daniel Fekete, and students.

7.5.1.1. Internships

- Jung Nam from the University of Minnesota visited for 3 months in the summer of 2018. His work centered on the use of storytelling mechanisms to support and communicate results of the exploration 3D data. This collaboration is still ongoing.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events Organisation

8.1.1.1. General Chair, Scientific Chair

- Jean-Daniel Fekete co-organized the Dagstuhl Seminar on Progressive Data Analysis and Visualization in 2018
- Jean-Daniel Fekete co-organized the Dagstuhl Seminar on Provenance and Logging for Sense Making in 2018

- Pierre Dragicevic co-organized the Journée Visu 2018.
- Tobias Isenberg co-organized for BELIV 2018.

8.1.2. Scientific Events Selection

8.1.2.1. Chair of Conference Program Committees

- Petra Isenberg was paper chair for InfoVis 2019.

8.1.2.2. Member of the Conference Program Committees

- Jean-Daniel Fekete was a member of the program committee for VIS 2018
- Jean-Daniel Fekete was a member of the best paper committee for PacificVis 2018.
- Pierre Dragicevic was a member of the program committee for VIS 2018.
- Tobias Isenberg was a member of the program committee for ACM/Eurographics Expressive 2018.
- Tobias Isenberg was a member of the program committee for IEEE VR 2018.
- Tobias Isenberg was a member of the program committee for EuroVis 2018.
- Tobias Isenberg was a member of the program committee for IVAPP 2018.
- Petra Isenberg was a member of the program committee for CHI.

8.1.2.3. Reviewer

- Jean-Daniel Fekete reviewed for CHI and PacificVis
- Pierre Dragicevic reviewed for VIS, UIST, EuroVis, alt.CHI, EICS, TEI, IHM.
- Tobias Isenberg reviewed for CHI, Expressive, InfoVis, PacificVis, SciVis, VAST, VISAP, and VR.
- Xiyao Wang reviewed for IHM, and TEI.
- Petra Isenberg reviewed for: CHI, EuroVA, EuroVis, ISS

8.1.3. Journal

8.1.3.1. Member of the Editorial Boards

- Pierre Dragicevic is member of the editorial board of the Journal of Perceptual Imaging (JPI).
- Pierre Dragicevic is member of the editorial board of the Springer Human–Computer Interaction Series (HCIS).
- Tobias Isenberg is member of the editorial board of Elsevier's Computers & Graphics journal.
- Petra Isenberg is member of the editorial board of IEEE Transactions on Visualization and Computer Graphics.
- Petra Isenberg is Associate Editor in Chief for IEEE Computer Graphics and Applications.

8.1.3.2. Reviewer - Reviewing Activities

- Jean-Daniel Fekete reviewed for TVCG
- Pierre Dragicevic reviewed for TOCHI, TVCG, JPI.
- Steve Haroz reviewed for Cognition, Journal of Experimental Psychology: General, Meta-Psychology
- Petra Isenberg reviewed for: Morgan Claypool (book proposal)

8.1.4. Invited Talks

- Jean-Daniel Fekete: Seminar EDF Lab Chatou "Progressive data analysis: a new language paradigm for scalability in exploratory data analysis", Chatou, Sep. 19, 2018
- Jean-Daniel Fekete: seminaire de recherche Medialab, « Nouvelles visualisations de réseaux pour de nouvelles formes d'explorations et d'analyses », Sciences Po Paris, April 3rd, 2018

- Jean-Daniel Fekete: Delft Data Science Seminar: Visual Data Science and its role in Computational Medicine The role of visualization in the hypothetico-deductive method, Delft University of Technology, January 6th, 2018.
- Tobias Isenberg: “Illustrative Visualization and Abstraction of Scientific Data”. M3DISIM joined Inria/Ecole Polytechnique research team, Orsay, France, November 2018.
- Tobias Isenberg: “Illustrative Visualization and Abstraction of Scientific Data”. GraphDeco research team, Sophia-Antipolis, France, December 2018.
- Steve Haroz: “Open Practices in Visualization Research”. Panel discussion at Evaluation and Beyond - Methodological Approaches for Visualization workshop, 21 Oct 2018.
- Petra Isenberg: “Evaluation in Visualization”. Keynote at the EMBL-EBI Workshop on Innovations in data visualisation for drug discovery, June 2018.
- Petra Isenberg: “Micro Visualizations for Pervasive and Mobile Data Exploration”. Seminar Series Talk, Technical University of Dresden, May 2018.

8.1.5. Leadership within the Scientific Community

- Jean-Daniel Fekete is the Chair of the Steering Committee of the IEEE Information Visualization Conference.
- Jean-Daniel Fekete is member of the VIS Executive Committee.
- Jean-Daniel Fekete is a member of the Eurographics Publication Board.
- Jean-Daniel Fekete is the Chair of the EuroVis Best Phd Award.
- Tobias Isenberg has been a member of the Executive Committee of the Visualization and Computer Graphics Technical Committee of the IEEE Computer Society and served as Publications Chair.
- Tobias Isenberg is a member of the Steering Committee of Expressive (Joined Symposium on Computational Aesthetics, Sketch-Based Interfaces & Modeling, and Non-Photorealistic Animation & Rendering).
- Steve Haroz is the Open Practices chair of the IEEE Information Visualization conference.

8.1.6. Scientific Expertise

- Pierre Dragicevic reviewed for an NSERC Discovery Grant proposal.

8.1.7. Research Administration

- Pierre Dragicevic: Member of the Commission Consultative de Spécialistes de l’Université Paris-Sud (CCSU).

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

- “Visualization and Visual Analytics” taught by Jean-Daniel Fekete, Master 2 in Data Science, 32 hours, École Polytechnique, France.
- “Photorealistic Rendering” taught by Tobias Isenberg at Polytech Paris-Sud and Université Paris-Saclay, France.
- “Introduction to Computer Graphics” taught by Tobias Isenberg at Polytech Paris-Sud, France.
- “Non-Photorealistic Rendering” taught by Tobias Isenberg at the University of Granada, Spain.
- guest lecture on “Introduction to Data Visualization” taught by Tobias Isenberg at Université Côte d’Azur, Nice, France.
- guest lecture on “Illustrative Visualization” taught by Tobias Isenberg at ENSTA, France.
- “Interactive Information Visualization” taught by Petra Isenberg at Université Paris Sud.
- “Visual Analytics” taught by Petra Isenberg at CentraleSupelec.

8.2.2. Supervision

- PhD: Marc Barnabé, Multiscale reconstruction of microbial ecosystems using semi-supervised machine learning, Université Paris-Sud, 2018, Jean-Daniel Fekete, Evelyne Lutton, INRA.
- PhD in progress: Mickaël Sereno, Collaborative Data Exploration and Discussion Supported by AR, Univ. Paris-Sud; 2018, Tobias Isenberg
- PhD in progress: Xiyao Wang, Augmented Reality Environments for the Interactive Exploration of 3D Data, Univ. Paris-Sud; 2017, Tobias Isenberg
- PhD in progress: Sarkis Halladjian, Spatially Integrated Abstraction of Genetic Molecules, Univ. Paris-Sud; 2017, Tobias Isenberg
- PhD in progress: Haichao Miao, Visual Abstraction and Modeling for DNA Nanotechnology, TU Wien, Austria, 2016, Tobias Isenberg

8.2.3. Juries

- Jean-Daniel Fekete: Member of the PhD committee of Dr. Marion Dumont.
- Jean-Daniel Fekete: Member of the PhD committee of Dr. Germán Leiva.
- Jean-Daniel Fekete: Member of the HdR committee of Dr. Eric Lecolinet.
- Pierre Dragicevic: Member of the hiring committee for Concours CRCN 2018 Centre Inria Rennes Bretagne Atlantique.
- Pierre Dragicevic: PhD defense committee of Bruno Fruchard (Télécom Paris Tech).
- Pierre Dragicevic: Mid-term PhD evaluation committee of Stacy Hsueh (Université Paris-Saclay).
- Pierre Dragicevic: Reviewer for Marguerite Peron's M2 internship.
- Petra Isenberg: Member of the PhD committee of Dr. Ulrike Kister
- Petra Isenberg: Reviewer Master Pro thesis, Ayoub Jaa, Université Paris Sud
- Petra Isenberg: Co-Supervisor Master thesis, Mathieu Louvet, Unversité Paris Sud
- Petra Isenberg: Co-Supervisor Master thesis, Mina Alipour, CentraleSupelec
- Petra Isenberg: Academic Tutor Master thesis: Marvin Rea

8.3. Popularization

8.3.1. Articles and contents

- Lonni Besançon's PhD work was featured in Le Monde's Binaire blog.
- Steve Haroz was interviewed on the Data Stories podcast, October 10, 2018

8.3.2. Creation of media or tools for science outreach

- Pierre Dragicevic and Yvonne Jansen: the data physicalization wiki and the List of Physical Visualizations and Related Artefacts (500 weekly visits) are continuously being updated.

9. Bibliography

Major publications by the team in recent years

[1] B. BACH, P. DRAGICEVIC, D. ARCHAMBAULT, C. HURTER, S. CARPENDALE.*A Review of Temporal Data Visualizations Based on Space-Time Cube Operations*, in "Eurographics Conference on Visualization", Swansea, Wales, United Kingdom, June 2014, https://hal.inria.fr/hal-01006140

[2] M. BEHRISCH, B. BACH, N. HENRY RICHE, T. SCHRECK, J.-D. FEKETE.*Matrix Reordering Methods for Table and Network Visualization*, in "Computer Graphics Forum", 2016, vol. 35, 24 [*DOI :* 10.1111/CGF.12935], https://hal.inria.fr/hal-01326759

[3] J. BOY, L. EVEILLARD, F. DETIENNE, J.-D. FEKETE.*Suggested Interactivity: Seeking Perceived Affordances for Information Visualization*, in "IEEE Transactions on Visualization and Computer Graphics", January 2016, vol. 22, n° 1, p. 639-648 [*DOI :* 10.1109/TVCG.2015.2467201], https://hal.inria.fr/hal-01188973

[4] S. BRUCKNER, T. ISENBERG, T. ROPINSKI, A. WIEBEL.*A Model of Spatial Directness in Interactive Visualization*, in "IEEE Transactions on Visualization and Computer Graphics", 2018, To appear [*DOI :* 10.1109/TVCG.2018.2848906], https://hal.inria.fr/hal-01813889

[5] E. DIMARA, A. BEZERIANOS, P. DRAGICEVIC.*The Attraction Effect in Information Visualization*, in "IEEE Transactions on Visualization and Computer Graphics", 2017, vol. 23, n° 1 [*DOI :* 10.1109/TVCG.2016.2598594], https://hal.inria.fr/hal-01355750

[6] P. ISENBERG, N. ELMQVIST, J. SCHOLTZ, D. CERNEA, K.-L. MA, H. HAGEN.*Collaborative Visualization: Definition, Challenges, and Research Agenda*, in "Information Visualization Journal (IVS)", 2011, vol. 10, n° 4, p. 310–326, Published online before print July 29, 2011 [*DOI :* 10.1177/1473871611412817], https://hal.inria.fr/inria-00638540

[7] T. ISENBERG, P. ISENBERG, J. CHEN, M. SEDLMAIR, T. MÖLLER.*A Systematic Review on the Practice of Evaluating Visualization*, in "IEEE Transactions on Visualization and Computer Graphics", 2013, vol. 19, n° 12, p. 2818-2827 [*DOI :* 10.1109/TVCG.2013.126], https://hal.inria.fr/hal-00846775

[8] Y. JANSEN, P. DRAGICEVIC, P. ISENBERG, J. ALEXANDER, A. KARNIK, J. KILDAL, S. SUBRAMANIAN, K. HORNBÆK.*Opportunities and Challenges for Data Physicalization*, in "Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)", New York, NY, United States, ACM, April 2015 [*DOI :* 10.1145/2702123.2702180], https://hal.inria.fr/hal-01120152

[9] H. LAM, E. BERTINI, P. ISENBERG, C. PLAISANT, S. CARPENDALE.*Empirical Studies in Information Visualization: Seven Scenarios*, in "IEEE Transactions on Visualization and Computer Graphics", 2012, vol. 18, n° 9, p. 1520–1536 [*DOI :* 10.1109/TVCG.2011.279], http://hal.inria.fr/hal-00932606

[10] M. LE GOC, L. H. KIM, A. PARSAEI, J.-D. FEKETE, P. DRAGICEVIC, S. FOLLMER.*Zooids: Building Blocks for Swarm User Interfaces*, in "Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST)", Tokyo, Japan, October 2016 [*DOI :* 10.1145/2984511.2984547], https://hal.inria.fr/hal-01391281

[11] B. LEE, P. ISENBERG, N. HENRY RICHE, S. CARPENDALE.*Beyond Mouse and Keyboard: Expanding Design Considerations for Information Visualization Interactions*, in "IEEE Transactions on Visualization and Computer Graphics", 2012, vol. 18, n° 12, p. 2689–2698 [*DOI :* 10.1109/TVCG.2012.204], https://hal.inria.fr/hal-00717203

[12] H. MIAO, E. DE LLANO, T. ISENBERG, M. E. GRÖLLER, I. BARIŠIC, I. VIOLA.*DimSUM: Dimension and Scale Unifying Maps for Visual Abstraction of DNA Origami Structures*, in "Computer Graphics Forum", June 2018, vol. 37, n° 3, p. 403-413 [*DOI :* 10.1111/CGF.13429], https://hal.inria.fr/hal-01795734

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

[14] N. PEZZOTTI, J.-D. FEKETE, T. HÖLLT, B. LELIEVELDT, E. EISEMANN, A. VILANOVA.*Multiscale Visualization and Exploration of Large Bipartite Graphs*, in "Computer Graphics Forum", June 2018, vol. 37, n^o 3, 12, https://hal.inria.fr/hal-01787046

[15] I. VIOLA, T. ISENBERG.*Pondering the Concept of Abstraction in (Illustrative) Visualization*, in "IEEE Transactions on Visualization and Computer Graphics", September 2018, vol. 24, n^o 9, p. 2573–2588 [*DOI :* 10.1109/TVCG.2017.2747545], https://hal.inria.fr/hal-01581177

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

Publications of the year

Articles in International Peer-Reviewed Journal

[17] T. BLASCHECK, L. BESANÇON, A. BEZERIANOS, B. LEE, P. ISENBERG.*Glanceable Visualization: Studies of Data Comparison Performance on Smartwatches*, in "IEEE Transactions on Visualization and Computer Graphics", January 2019, vol. 25, n^o 1, https://hal.inria.fr/hal-01851306

[18] N. BOUKHELIFA, M. BRYANT, N. BULATOVIĆ, I. ČUKIĆ, J.-D. FEKETE, M. KNEŽEVIĆ, J. LEHMANN, D. STUART, C. THIEL.*The CENDARI Infrastructure*, in "Journal on Computing and Cultural Heritage", April 2018, vol. 11, n^o 2, p. 1-20 [*DOI :* 10.1145/3092906], https://hal.inria.fr/hal-01523102

[19] M. BREHMER, B. LEE, P. ISENBERG, E. K. CHOE.*Visualizing Ranges over Time on Mobile Phones: A Task-Based Crowdsourced Evaluation*, in "IEEE Transactions on Visualization and Computer Graphics", August 2018, vol. 25, p. 1-11, https://hal.inria.fr/hal-01857469

[20] S. BRUCKNER, T. ISENBERG, T. ROPINSKI, A. WIEBEL.*A Model of Spatial Directness in Interactive Visualization*, in "IEEE Transactions on Visualization and Computer Graphics", 2018, To appear [*DOI :* 10.1109/TVCG.2018.2848906], https://hal.inria.fr/hal-01813889

[21] E. DIMARA, G. BAILLY, A. BEZERIANOS, S. FRANCONERI.*Mitigating the Attraction Effect with Visualizations*, in "IEEE Transactions on Visualization and Computer Graphics", October 2018, 11 [*DOI :* 10.1109/TVCG.2018.2865233], https://hal.inria.fr/hal-01845004

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

[23] E. DIMARA, S. FRANCONERI, C. PLAISANT, A. BEZERIANOS, P. DRAGICEVIC.*A Task-based Taxonomy of Cognitive Biases for Information Visualization*, in "IEEE Transactions on Visualization and Computer Graphics", 2019, https://hal.sorbonne-universite.fr/hal-01868738

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

[25] J. JO, J. SEO, J.-D. FEKETE.*PANENE: A Progressive Algorithm for Indexing and Querying Approximate k-Nearest Neighbors*, in "IEEE Transactions on Visualization and Computer Graphics", 2018 [*DOI :* 10.1109/TVCG.2018.2869149], https://hal.inria.fr/hal-01855672

[26] J. JO, F. VERNIER, P. DRAGICEVIC, J.-D. FEKETE.*A Declarative Rendering Model for Multiclass Density Maps*, in "IEEE Transactions on Visualization and Computer Graphics", January 2019, vol. 25, n^{o} 1, p. 470-480 [*DOI :* 10.1109/TVCG.2018.2865141], https://hal.inria.fr/hal-01848427

[27] K. LAWONN, I. VIOLA, B. PREIM, T. ISENBERG.*A Survey of Surface-Based Illustrative Rendering for Visualization*, in "Computer Graphics Forum", September 2018, vol. 37, n^{o} 6, p. 205–234 [*DOI :* 10.1111/CGF.13322], https://hal.inria.fr/hal-01652236

[28] M. LE GOC, C. PERIN, S. FOLLMER, J.-D. FEKETE, P. DRAGICEVIC.*Dynamic Composite Data Physicalization Using Wheeled Micro-Robots*, in "IEEE Transactions on Visualization and Computer Graphics", January 2019, vol. 25, n^{o} 1, p. 737-747 [*DOI :* 10.1109/TVCG.2018.2865159], https://hal.inria.fr/hal-01848436

[29] H. MIAO, E. DE LLANO, T. ISENBERG, M. E. GRÖLLER, I. BARIŠIC, I. VIOLA.*DimSUM: Dimension and Scale Unifying Maps for Visual Abstraction of DNA Origami Structures*, in "Computer Graphics Forum", June 2018, vol. 37, n^{o} 3, p. 403-413 [*DOI :* 10.1111/CGF.13429], https://hal.inria.fr/hal-01795734

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

[31] H. MIAO, T. KLEIN, D. KOUŘIL, P. MINDEK, K. SCHATZ, M. E. GRÖLLER, B. KOZLÍKOVÁ, T. ISENBERG, I. VIOLA.*Multiscale Molecular Visualization*, in "Journal of Molecular Biology", 2018, To appear [*DOI :* 10.1016/J.JMB.2018.09.004], https://hal.inria.fr/hal-01873724

[32] A. NEWMAN, Z. BYLINSKII, S. HAROZ, S. MADAN, F. DURAND, A. OLIVA.*Effects of title wording on memory of trends in line graphs*, in "Journal of Vision", 2018, vol. 18, n^{o} 10, 837, https://hal.inria.fr/hal-01947467

[33] C. PERIN, J.-D. FEKETE, P. DRAGICEVIC.*Jacques Bertin's Legacy in Information Visualization and the Reorderable Matrix*, in "Cartography and Geographic Information Science", May 2018 [*DOI :* 10.1080/15230406.2018.1470942], https://hal.inria.fr/hal-01786606

[34] N. PEZZOTTI, J.-D. FEKETE, T. HÖLLT, B. LELIEVELDT, E. EISEMANN, A. VILANOVA.*Multiscale Visualization and Exploration of Large Bipartite Graphs*, in "Computer Graphics Forum", June 2018, vol. 37, n^{o} 3, 12, https://hal.inria.fr/hal-01787046

[35] I. VIOLA, T. ISENBERG.*Pondering the Concept of Abstraction in (Illustrative) Visualization*, in "IEEE Transactions on Visualization and Computer Graphics", September 2018, vol. 24, n^{o} 9, p. 2573–2588 [*DOI :* 10.1109/TVCG.2017.2747545], https://hal.inria.fr/hal-01581177

International Conferences with Proceedings

[36] B. BACH, C. PERIN, Q. REN, P. DRAGICEVIC.*Ways of Visualizing Data on Curves*, in "TransImage 2018 - 5th Biennial Transdisciplinary Imaging Conference", Edinburgh, United Kingdom, April 2018, p. 1-14 [*DOI :* 10.6084/M9.FIGSHARE.6104705], https://hal.inria.fr/hal-01818137

[37] L. BATTLE, M. ANGELINI, C. BINNIG, T. CATARCI, P. EICHMANN, J.-D. FEKETE, G. SANTUCCI, M. SEDLMAIR, W. WILLETT.*Evaluating Visual Data Analysis Systems: A Discussion Report*, in "HILDA'18: Workshop on Human-In-the-Loop Data Analytics", Houston, TX, United States, June 2018 [*DOI :* 10.1145/3209900.3209901], https://hal.inria.fr/hal-01786507

[38] L. BESANÇON, A. SEMMO, D. J. BIAU, B. FRACHET, V. PINEAU, E. H. SARIALI, R. TAOUACHI, T. ISENBERG, P. DRAGICEVIC.*Reducing Affective Responses to Surgical Images through Color Manipulation and Stylization*, in "Proceedings of the Joint Symposium on Computational Aesthetics, Sketch-Based Interfaces and Modeling, and Non-Photorealistic Animation and Rendering", Victoria, Canada, ACM, August 2018, p. 4:1–4:13 [*DOI :* 10.1145/3229147.3229158], https://hal.inria.fr/hal-01795744

[39] T. BLASCHECK, A. BEZERIANOS, L. BESANÇON, B. LEE, P. ISENBERG.*Preparing for Perceptual Studies: Position and Orientation of Wrist-worn Smartwatches for Reading Tasks*, in "Proceedings of the Workshop on Data Visualization on Mobile Devices held at ACM CHI", Montréal, Canada, April 2018, https://hal.inria.fr/hal-01744246

[40] T. BLASCHECK, M. BURCH, T. MEISEL, T. SCHNEIDER, S. MUMIN.*Exploring Eye Movements with Node-Link Graph Layouts*, in "Eye Tracking for Spatial Research, Proceedings of the 3rd International Workshop", Zürich, Switzerland, P. KIEFER, I. GIANNOPOULOS, F. GÖBEL, M. RAUBAL, A. T. DUCHOWSKI (editors), January 2018 [*DOI :* 10.3929/ETHZ-B-000222279], https://hal.inria.fr/hal-01766812

[41] S. HAROZ.*Open Practices in Visualization Research*, in "Evaluation and Beyond - Methodological Approaches for Visualization", Berlin, Germany, October 2018 [*DOI :* 10.31219/OSF.IO/8AG3W], https://hal.inria.fr/hal-01947432

[42] R. KOSARA, S. HAROZ.*Skipping the Replication Crisis in Visualization: Threats to Study Validity and How to Address Them*, in "Evaluation and Beyond - Methodological Approaches for Visualization", Berlin, Germany, October 2018 [*DOI :* 10.31219/OSF.IO/F8QEY], https://hal.inria.fr/hal-01947436

[43] B. PREIM, T. ROPINSKI, P. ISENBERG.*A Critical Analysis of the Evaluation Practice in Medical Visualization*, in "Eurographics 8th EG Workshop on Visual Computing for Biology and Medicine", Granada, Spain, VCBM 2018, September 2018 [*DOI :* 10.2312/VCBM.20181228], https://hal.inria.fr/hal-01893153

[44] M. TREFZGER, T. BLASCHECK, M. RASCHKE, S. HAUSMANN, T. SCHLEGEL.*A Visual Comparison of Gaze Behavior from Pedestrians and Cyclists*, in "ETRA 2018 - Symposium on Eye Tracking Research and Applications", Warsaw, Poland, ACM, June 2018 [*DOI :* 10.1145/3204493.3204553], https://hal.inria.fr/hal-01766815

[45] C. WACHARAMANOTHAM, S. GUHA, M. KAY, P. DRAGICEVIC, S. HAROZ.*Special Interest Group on Transparent Statistics Guidelines*, in "CHI 2018 - Conference on Human Factors in Computing Systems", Montreal QC, France, ACM Press, April 2018, p. 1-4 [*DOI :* 10.1145/3170427.3185374], https://hal.inria.fr/hal-01818141

Conferences without Proceedings

[46] T. Blascheck, L. Besançon, A. Bezerianos, B. Lee, P. Isenberg.*Visualization perception on smartwatches*, in "Journée Visu 2018", Palaiseau, France, May 2018, https://hal.inria.fr/hal-01844726

[47] A. Holcombe, C. Ludowici, S. Haroz.*Is there a reproducibility crisis around here? Maybe not, but we still need to change*, in "Vision Sciences Society", St. Pete Beach, Florida, United States, 2019, https://hal.inria.fr/hal-01944513

[48] P. Valdivia, P. Buono, C. Plaisant, N. Dufournaud, J.-D. Fekete.*Using Dynamic Hypergraphs to Reveal the Evolution of the Business Network of a 17th Century French Woman Merchant*, in "VIS 2018 - 3rd Workshop on Visualization for the Digital Humanities", Berlin, Germany, October 2018, https://hal.inria.fr/hal-01855675

Scientific Books (or Scientific Book chapters)

[49] W. Büschel, J. Chen, R. Dachselt, S. M. Drucker, T. Dwyer, C. Görg, T. Isenberg, A. Kerren, C. North, W. Stuerzlinger.*Interaction for Immersive Analytics*, in "Immersive Analytics", Springer, 2018, p. 95 - 138 [*DOI :* 10.1007/978-3-030-01388-2_4], https://hal.inria.fr/hal-01907526

[50] T. Dwyer, K. Marriott, T. Isenberg, K. Klein, N. Riche, F. Schreiber, W. Stuerzlinger, B. Thomas.*Immersive Analytics: An Introduction*, in "Immersive Analytics", Springer, 2018, p. 1 - 23 [*DOI :* 10.1007/978-3-030-01388-2_1], https://hal.inria.fr/hal-01907533

[51] B. Thomas, G. Welch, P. Dragicevic, N. Elmqvist, P. Irani, Y. Jansen, D. Schmalstieg, A. Tabard, N. ElSayed, R. Smith, W. Willett.*Situated Analytics*, in "Immersive Analytics", Lecture Notes in Computer Science, Springer, October 2018, vol. 11190, p. 185-220 [*DOI :* 10.1007/978-3-030-01388-2_7], https://hal.inria.fr/hal-01947243

Books or Proceedings Editing

[52] E. Lutton, P. Legrand, P. Parrend, N. Monmarché, M. Schoenauer (editors). *Artificial Evolution, 13th International Conference, Evolution Artificielle, EA 2017, Paris, France, October 25-27, 2017,Revised Selected Papers*, Artificial Evolution, Springer, March 2018, vol. LNCS - Lecture Notes in Computer Science, n^{o} 10764 [*DOI :* 10.1007/978-3-319-78133-4], https://hal.inria.fr/hal-01910443

Research Reports

[53] R. Chang, J.-D. Fekete, J. Freire, C. Scheidegger.*Connecting Visualization and Data Management Research (Dagstuhl Seminar 17461)*, Dagstuhl, 2018, n^{o} 2018/8670, p. 46–58 [*DOI :* 10.4230/DagRep.7.11.46], https://hal.inria.fr/hal-01756799

Scientific Popularization

[54] X. Wang, L. Besançon, M. Ammi, T. Isenberg.*Navigation Tactile 3D Augmente´e Sur Mobiles*, in "Journée Visu 2018", Palaiseau, France, May 2018, https://hal.inria.fr/hal-01854463

Other Publications

[55] M. ALIPOUR, P. DRAGICEVIC, T. ISENBERG, P. ISENBERG.*Situated Visualizations of Office Noise to Promote Personal Health*, October 2018, p. 1-2, InfoVis 2018 - IEEE Conference on Information Visualization, Poster, https://hal.inria.fr/hal-01857354

[56] E. DE LLANO, H. MIAO, T. ISENBERG, M. E. GRÖLLER, I. VIOLA, I. BARIŠIC.*A Preview to Adenita: Modeling and Visualization of DNA Nanostructures*, June 2018, Posters at the 3rd Functional DNA Nanotechnology Workshop, Poster, https://hal.inria.fr/hal-01813901

[57] P. ISENBERG, C. KINKELDEY, J.-D. FEKETE.*Visual Analytics for Monitoring and Exploration of Blockchain Data With a Focus on the Bitcoin Blockchain*, 2018, HCI for Blockchain: A CHI 2018 workshop on Studying, Critiquing, Designing and Envisioning Distributed Ledger Technologies, https://hal.inria.fr/hal-01950934

[58] C. TURKAY, N. PEZZOTTI, C. BINNIG, H. STROBELT, B. HAMMER, D. A. KEIM, J.-D. FEKETE, T. PALPANAS, Y. WANG, F. RUSU.*Progressive Data Science: Potential and Challenges*, December 2018, https://arxiv.org/abs/1812.08032 - working paper or preprint, https://hal.inria.fr/hal-01961871

References in notes

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

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

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

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

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

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

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

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

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

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

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

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

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

Project-Team CEDAR

Rich Data Exploration at Cloud Scale

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

IN PARTNERSHIP WITH:
Ecole Polytechnique

RESEARCH CENTER
Saclay - Île-de-France

THEME
Data and Knowledge Representation and Processing

Table of contents

Project-Team CEDAR

Creation of the Team: 2016 January 01, updated into Project-Team: 2018 April 01

Keywords:

Computer Science and Digital Science:

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

Other Research Topics and Application Domains:

B8.5.1. - Participative democracy
B9.5.6. - Data science
B9.7.2. - Open data

1. Team, Visitors, External Collaborators

Research Scientists

Ioana Manolescu [Team leader, Inria, Senior Researcher, HDR]
Yanlei Diao [Ecole Polytechnique, Senior Researcher, HDR]
Michael Thomazo [Inria, Researcher, until Mar 2018]

External Collaborators

Pawel Guzewicz [Université Paris Saclay, until Feb 2018]
Lars Kegel [Université de Dresden, until Aug 2018]
Minh Huong Le Nguyen [Ecole polytechnique, from Sep 2018]
Arnaud Stiegler [Ecole polytechnique, from May 2018 until Aug 2018]
Xavier Tannier [CNRS]
Stamatios Zampetakis [Orchestra Networks]

Technical Staff

Ahmed Abdelkafi [Ecole polytechnique, until Jul 2018]
Laurent Cetinsoy [Ecole polytechnique, from Oct 2018]
Tayeb Merabti [Inria]

PhD Students

Maxime Buron [Inria]
Tien Duc Cao [Inria]
Sejla Cebiric [Inria, until Jul 2018]
Luciano Di Palma [Ecole polytechnique]
Pawel Guzewicz [Ecole polytechnique, from Oct 2018]
Enhui Huang [Ecole polytechnique]
Felix Raimundo [Ecole polytechnique]
Alexandre Sevin [Ecole polytechnique, until Apr 2018]
Khaled Zaouk [Ecole polytechnique]

Post-Doctoral Fellows

Mirjana Mazuran [Inria, from Jul 2018]
Fei Song [Inria]

Visiting Scientists

Hugo Cisneros [CNRS, from Apr 2018 until Aug 2018]
Juliana Freire [Digiteo, from Aug 2018]
Minh Huong Le Nguyen [Ecole polytechnique, until Feb 2018]

Administrative Assistant

Maeva Jeannot [Inria]

2. Overall Objectives

2.1. Overall Objectives

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

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

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

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

3. Research Program

3.1. Scalable Heterogeneous Stores

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

3.2. Semantic Query Answering

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

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

3.3. Multi-Model Querying

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

3.4. Interactive Data Exploration at Scale

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

3.5. Exploratory Querying of Semantic Graphs

Semantic graphs including data and knowledge are hard to apprehend for users, due to the complexity of their structure and oftentimes to their large volumes. To help tame this complexity, in prior research (2014), we have presented a full framework for RDF data warehousing, specifically designed for heterogeneous and semantic-rich graphs. However, this framework still leaves to the users the burden of chosing the most interesting warehousing queries to ask. More user-friendly data management tools are needed, which help the user discover the interesting structure and information hidden within RDF graphs. This research has benefitted from the arrival in the team of Mirjana Mazuran, as well as from the start of the PhD thesis of Paweł Guzewicz, co-advised by Yanlei Diao and Ioana Manolescu.

4. Application Domains

4.1. Cloud Computing

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

4.2. Computational Journalism

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

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

4.3. Open Data Intelligence

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

4.4. Genomics

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

5. Highlights of the Year

5.1. Highlights of the Year

Conference Chair

Ioana Manolescu has been a general chair of the IEEE International Conference on Data Engineering (ICDE) 2018.

Keynotes

Ioana Manolescu has given invited keynote talks at the Extended Semantic Web Conference (ESWC) 2018 [25], and at the *34ème Conférence sur la Gestion de Données – Principes, Technologies et Applications* (BDA) 2018 [24].

PVLDB paper

A paper on "Optimization for active learning-based interactive database exploration" by Enhui Huang and co-authors has been accepted at PVLDB 2018 [10].

Prix de stage de l'Ecole Polytechnique

Camille Chanial, third-year (M1) student at Ecole Polytechnique, has been awarded a Prix de Stage for his work on the ConnectionLens prototype [9].

6. New Software and Platforms

6.1. Tatooine

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

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

6.2. AIDES

KEYWORDS: Data Exploration - Active Learning
FUNCTIONAL DESCRIPTION: AIDES is a data exploration software. It allows a user to explore a huge (tabular) dataset and discover tuples matching his or her interest. Our system repeatedly proposes the most informative tuples to the user, who must annotate them as "interesting" / "not-interesting", and as iterations progress an increasingly accurate model of the user's interest region is built. Our system also focuses on supporting low selectivity, high-dimensional interest regions.

- Contact: Yanlei Diao

6.3. OntoSQL

KEYWORDS: RDF - Semantic Web - Querying - Databases

FUNCTIONAL DESCRIPTION: OntoSQL is a tool providing three main functionalities: - Loading RDF graphs (consisting of data triples and possibly a schema or ontology) into a relational database, - Saturating the data based on the ontology. Currently, RDF Schema ontologies are supported. - Querying the loaded data using conjunctive queries. Data can be loaded either from distinct files or from a single file containing them both. The loading process allows to choose between two storage schemas: - One triples table. - One table per role and concept. Querying provides an SQL translation for each conjunctive query according to the storage schema used in the loading process, then the SQL query is evaluated by the underlying relational database.

- Participants: Ioana Manolescu, Michaël Thomazo and Tayeb Merabti
- Partner: Université de Rennes 1
- Contact: Ioana Manolescu
- URL: https://ontosql.inria.fr/

6.4. ConnectionLens

KEYWORDS: Data management - Big data - Information extraction - Semantic Web
FUNCTIONAL DESCRIPTION: ConnectionLens treats a set of heterogeneous, independently authored data sources as a single virtual graph, whereas nodes represent fine-granularity data items (relational tuples, attributes, key-value pairs, RDF, JSON or XML nodes. . .) and edges correspond either to structural connections (e.g., a tuple is in a database, an attribute is in a tuple, a JSON node has a parent. . .) or to similarity (sameAs) links. To further enrich the content journalists work with, we also apply entity extraction which enables to detect the people, organizations etc. mentioned in text, whether full-text or text snippets found e.g. in RDF or XML. ConnectionLens is thus capable of finding and exploiting connections present across heterogeneous data sources without requiring the user to specify any join predicate.

- Contact: Manolescu Ioana
- Publication: ConnectionLens: Finding Connections Across Heterogeneous Data Sources

6.5. INSEE-Extract

Spreadsheets extractor
KEYWORDS: RDF - Data extraction
FUNCTIONAL DESCRIPTION: Extract content of spreadsheets automatically and store it as RDF triples

- Participants: Ioana Manolescu, Xavier Tannier and Tien Duc Cao
- Contact: Tien Duc Cao
- Publication: Extracting Linked Data from statistic spreadsheets
- URL: https://gitlab.inria.fr/cedar/excel-extractor

6.6. INSEE-Search

KEYWORDS: Document ranking - RDF
FUNCTIONAL DESCRIPTION: Searching for relevant data cells (or data row/column) given a query in natural language (French)

- Participants: Ioana Manolescu, Xavier Tannier and Tien Duc Cao
- Contact: Tien Duc Cao
- Publications: Extracting Linked Data from statistic spreadsheets - Searching for Truth in a Database of Statistics

6.7. RDFQuotient

Quotient summaries of RDF graphs
KEYWORDS: RDF - Graph algorithmics - Graph visualization - Graph summaries - Semantic Web

FUNCTIONAL DESCRIPTION: RDF graphs can be large and heterogeneous, making it hard for users to get acquainted with a new graph and understand whether it may have interesting information. To help users figure it out, we have devised novel equivalence relations among RDF nodes, capable of recognizing them as equivalent (and thus, summarize them together) despite the heterogeneity often exhibited by their incoming and outgoing node properties. From these relations, we derive four novel summaries, called Weak, Strong, Typed Weak and Typed Strong, and show how to obtain from them compact and enticing visualizations.

- Partner: Université de Rennes 1
- Contact: Manolescu Ioana
- Publications: Compact Summaries of Rich Heterogeneous Graphs - Structural Summarization of Semantic Graphs

7. New Results

7.1. Interactive Data Exploration at Scale

Building upon our prior work in active learning-based interactive database exploration system, we improved this system in terms of efficiency and effectiveness. First, we formally defined the class of user interest queries to which our proposed Dual Space Model (DSM) can bring significant improvement in accuracy. Second, we generalized the DSM to arbitrary queries by forcing our system to fall back to the traditional active learning-based techniques if the requested query properties are not satisfied. Third, we launched a user study to collect real-world datasets and user interest patterns for comparison experiments. The evaluation results showed that our new system outperformed the start-of-the-art active learning techniques and data exploration systems. Fourth, to show the robustness of our system, we added some label noise into the experiments. It turned out that our system maintained a good performance and significantly outperformed traditional active learning-based system. These results have appeared in the prestigious PVLDB journal [10]. In addition, we have been working on integrating DSM with version space algorithms and designing more advanced methods to deal with label noise. In the near feature, a new software based on our proposed techniques will be put into use for interactive database exploration.

7.2. A learning-based approach to optimizing large-scale data analytics

As part of my PhD thesis of K. Zaouk, we have proposed two neural network architectures to support in-situ modeling of user objectives in large-scale data analytics. Although conceptually these architectures can work with any big data system, the modeling of user-objectives on analytics run was applied on Spark Streaming. In our problem settings where only few traces are run whenever a new workload is submitted to the cloud, we have proposed new optimizations to improve the accuracy and efficiency of the auto-encoder based architecture. Thus, we have developed a prototype that included these neural network architectures and optimizations. This prototype was then used to evaluate a benchmark of stream analytics that we developed and instrumented on top of two clusters that collect Spark Streaming workloads' traces.

We analyzed the performance of the proposed techniques and demonstrated their performance benefits over state of the art performance modeling techniques based on machine learning (such as Ottertune used in tuning traditional RDBMS). Our latest results show that we outperform Ottertune in robustness and in our problem settings. These results consolidated in a paper "Boosting Big Data Analytics with Deep Learning Models and Optimization Methods" submitted for publication, alongside with other scientific results in multi-objective optimization contributed by the co-author Fei Song. Work on this topic continues.

7.3. Event stream analysis

As enterprise information systems are collecting event streams from various sources, the ability of a system to automatically detect anomalous events and further provide human readable explanations is of paramount importance. In a position paper [19], we argue for the need of a new type of data stream analytics that can address anomaly detection and explanation discovery in a single, integrated system, which not only offers increased business intelligence, but also opens up opportunities for improved solutions. In particular, we propose a two-pass approach to building such a system, highlight the challenges, and offer initial directions for solutions.

7.4. Quotient summarization of RDF graphs

We have continued our work on efficiently computing informative summaries of large, heterogeneous RDF graphs.

First, we have noticed that type information, when available, can be used to group RDF nodes in interesting, pertinent equivalence classes. However, the integration of type in our quotient summarization framework (presented in ISWC 2017) is not straightforward, since an RDF node may have zero, one, or more than one types. In [15], we have identified a sufficient, flexible condition under which we are able to propose a form of quotient summarization based on types, even if a node has multiple types, and even if they are not organized in a tree-shape classification, but instead in a directed acyclic graph (DAG).

In parallel, we have finalized a comprehensive survey of RDF graph summarization techniques which appeared in the VLDB Journal [8]. We have also completely re-developed our RDF graph summarization platform, in order to ensure correctness, to factorize common elements across all the summarization methods, and to implement new, incremental summarization algorithms [21]. This work has attracted significant visibility through an invited keynote at the ESWC conference [25], and through an ISWC "Resource" publication where our summaries are integrated in a LOD visual exploration portal developed by the ILDA team of Inria [17].

7.5. Semantic integration of heterogeneous data

A large amount of data sources are publicly available in *heterogeneous formats* such as relational, RDF and JSON. These data sources can share information about common entities, which the users may want to query as a single dataset, possibly exploiting also a set of semantic constraints which serve as a common integration perspective. We proposed a new approach to query such *integration* of datasources in a *global RDF graph* using an *RDFS ontology* and user-specified entailment rules. Previous approaches to query answering in the presence of knowledge involve either the materialization of inferred data, or reformulation of the query; both approches have well-known drawbacks. We introduce a new way of query answering as a reduction to view-based answering in [11]. This approach avoids both materialization in the data and query reformulation.

We have also developed an RDF Schema *reformulation algorithm* taking into account the reasoning on the ontology. This algorithm reduces query answering on data in the presence of an ontology, to query evaluation (solely on the data). In particular, this reformulation algorithm can be used to speed up query answering in the integration system mentioned above.

7.6. Fact-checking: a content management perspective

Throughout the year, we have worked within the ANR ContentCheck project to analyze and systematize computational fact-checking as a discipline of computer science; we have analyzed and classified existing works in this area, proposed a generic architecture for computational fact-checking, and highlighted perspectives in a Web Conference (formerly known as WWW) article [18] and two tutorials, presented respectively at the Web conference [16] and the PVLDB conference [7]. This work has also been featured in an invited keynote at the BDA 2018 conference [24].

7.7. Novel fact-checking architectures and algorithms

Still part of our work in ContentCheck, we have worked to devise new algorithms and architectures for data journalism and journalistic fact checking.

First, we have considered the problem of making it easy to check the accuracy of a statistic claim, in the statistic database published by INSEE, the leading french statistic institute. In prior work, we had shown how the INSEE data can be converted into a collection of open data adherent to the best practices of the W3C (RDF graphs). Following up on that work, we have proposed a novel algorithm which allows to search these RDF datasets by means of user-friendly keyword queries. Our algorithm returns ranked answers at the granularity of the RDF dataset (corresponding to a spreadsheet in a statistic dataset published by INSEE) or, when possible, at the granularity of individual cells, or line/column in a spreadsheet that best matches the user query [13], [12].

Second, we have devised a new architecture for keyword search in a polystore systems, where users ask a set of keywords, and receive results showing how occurrences of these keywords across the set of data sources can be connected. This allows identifying possibly unforeseen connections across heterogeneous data sources. We have implemented this architecture in the ConnectionLens prototype, which we demonstrated in VLDB [9] and also informally at BDA [14].

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

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

8.1.2. LabEx, IdEx

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

8.1.3. Others

- ODIN is a four-year project started (2014-2018) funded by the Direction Générale de l’Armement, between the SemSoft company, IRISA Rennes and Cedar. The project focused on developing a complete framework for analytics on Web data, in particular taking into account uncertainty, based on Semantic Web technologies such as RDF.

- The goal of the iCODA project is to develop the scientific and technological foundations for knowledge- mediated user-in-the-loop collaborative data analytics on heterogenous information sources, and to demonstrate the effectiveness of the approach in realistic, high-visibility use-cases. The project stands at the crossroad of multiple research fields—content analysis, data management, knowledge represen- tation, visualization—that span multiple Inria themes, and counts on a club of major press partners to define usage scenarios, provide data and demonstrate achievements. This is a project funded directly by Inria ("Inria Project Lab"), and is in collaboration with GraphIK, ILDA, LINKMEDIA (coordinator), as well as the press partners AFP, Le Monde (Les Décodeurs) and Ouest-France.

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

- **IDEAA: Issue-Driven European Arena Analytics** is a project funded by the European Commission Union's Horizon 2020 research and innovation programme. The project started in July 2018 for a duration of two years. Its purpose is to allow citizens to easily explore the trove of publicly available data with the aim of building a viewpoint on specific issues. Its main strengths are: supply users with succinct and meaningful knowledge with respect to the issue they are interested in; allow users to interact with the provided knowledge to refine their information need and advance understanding; suggest interesting or unexpected aspects in the data and support the comparison of knowledge discovered from different data sources. IDEAA is inspired by human-to-human dialogues, where questions are explorative, possibly imprecise, and answers may be a bit inaccurate but suggestive, conveying an idea that stimulates the interlocutor to further questions.

  The project supports a two-years presence of Mirjana Mazuran as an experienced post-doc in our team.

8.3. International Initiatives

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

8.3.1.1. WebClaimExplain

Title: Mining for explanations to claims published on the Web

International Partner (Institution - Laboratory - Researcher):

AIST (Japan) - Julien Leblay

Start year: 2017

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

The goal of this research is to create tools to find explanations for facts and verify claims made online. While this process cannot be fully automated, the main focus of our work will be explanation finding via trusted sources, based on the observation that one can only trust a statement if he/she can explain it through rules and proofs that can themselves be trusted. Our WebClaimExplain collaboration has been particularly fruitful this year in terms of publications [9], [7], [18], [16], [14].

8.3.2. Inria International Partners

8.3.2.1. Informal International Partners

We resumed our collaboration with Prof. Alin Deutsch from University of California in San Diego (UCSD), during his invited stay at U. Paris Sud. We have completed a work (started in 2015-2016) on efficient view-based query rewriting in polystores, and submitted it to a major international conference.

8.3.3. Participation in Other International Programs

8.3.3.1. AYAME

WebClaimExplain

Title: Mining for explanations to claims published on the Web

International Partner (Institution - Laboratory - Researcher):

AIST (Japan) - Leblay Julien

Duration: 2017 - 2019

Start year: 2017

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

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

8.4. International Research Visitors

8.4.1. Visits of International Scientists

8.4.1.1. Sabbatical programme

Juliana Freire, a professor at NYU and the chair of the ACM SIGMOD chapter, has been a visitor on sabbatical in the team since September 2018.

8.4.1.2. Internships

Lars Kegel, a PhD student at the university of Dresden, has visited the team until August 2018. He has worked on characterizing and generating time series data for benchmarking time series management software.

8.4.2. Visits to International Teams

8.4.2.1. Research Stays Abroad

Yanlei Diao spent three months at U. Massachussets at Amherst, USA.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

Ioana Manolescu has been a general chair of the IEEE International Conference on Data Engineering (ICDE) 2018.

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

Ioana Manolescu has been a member of the program committees of: the International Workshop on Semantic Big Data (SBD), in cojunction with ACM SIGMOD 2018; the Data Engineering Meets Semantic Web Workshop (DESWeb) in conjunction with IEEE ICDE 2018; the WebDB workshop, in conjunction with ACM SIGMOD 2018; the 34th *Conférence sur la Gestion de Données – Principes, Technologies et Applications* (BDA) 2018.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

Yanlei Diao has been the Editor-in-Chief of the ACM SIGMOD Record, SIGMOD's quarterly newsletter.

Ioana Manolescu has been an associate editor of the PVLDB (Proceedings of VLDB) Journal 2018.

9.1.4. Invited Talks

Ioana Manolescu has given invited keynotes at the Extended Semantic Web Conference (ESWC) 2018 , and at the *34ème Conférence sur la Gestion de Données – Principes, Technologies et Applications* (BDA) 2018.

9.1.5. Leadership within the Scientific Community

Yanlei Diao and Ioana Manolescu are both members of the PVLDB Endowment Board of Trustees.

Ioana Manolescu has been the president the ACM SIGMOD "Jim Gray" PhD Award Committee.

Ioana Manolescu is a member of the steering committee (*Comité de Pilotage*) of "Bases de Données Avancées" (BDA), the informal association organizing the database research community in France and french-speaking countries.

9.1.6. Research Administration

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

Ioana Manolescu has been a member of Inria's Commission d'Evaluation and of the Bureau des Comités de Projets of Inria Saclay.

Team members have participated to the following hiring committees:

- Associate professor – Ecole Polytechnique, Yanlei Diao. Jury member.
- Full professor – INSA Lyon, Ioana Manolescu. Jury member.
- Associate professor – Ecole Polytechnique, Ioana Manolescu. Jury member.
- Inria researcher – Inria Nancy, Ioana Manolescu. Jury member.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

- Master: Y. Diao is a Professor at Ecole Polytechnique, where she teaches "System for Big Data" in M1; she also teaches "Systems for Big Data Analytics" in M2 in the Data Science Master Program of Université Paris Saclay.
- Master: I. Manolescu, Architectures for Massive Data Management, 12h, M2, Université Paris-Saclay.
- Master: I. Manolescu, Database Management Systems, 52h, M1, École Polytechnique.
- Licence: I. Manolescu, Giant Global Graph, 18h, L3, École Polytechnique.

9.2.2. Supervision

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

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

PhD interrupted in July 2018: Sejla Čebirić: "CloudSelect: Data Mobility Within, Across and Outside Clouds", since September 2015, François Goasdoué Goasdoué and Ioana Manolescu. The student joined the industry.

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

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

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

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

PhD interrupted in April 2018: Alexandre Sevin: "Exploration interactive de données sur de grandes sources de données hétérogènes", since October 2017, Yanlei Diao and Peter Haas (U. Massachussets at Amherst, USA). The student joined the industry.

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

9.2.3. Juries

Ioana Manolescu reported on the PhD thesis of Louis Jachiet (Inria Grenoble).

9.3. Popularization

9.3.1. Articles and contents

- In books/journals for the general public: Our research on computational fact-checking has been featured in Le Journal Toulousain in an article titled "La science se met au service de l'information" , in February 2018 (http://www.lejournaltoulousain.fr/societe/la-science-se-met-au-service-de-linformation-pour-detecter-les-fake-news-54906). Further, I. Manolescu has been interviewed as an expert for an article in Le Figaro, in March 2018 (http://www.lefigaro.fr/secteur/high-tech/2018/03/08/32001-20180308ARTFIG00341-sur-twitter-les-fake-news-voyagent-plus-vite-que-les-vraies-informations.php).

9.3.2. Interventions

- Ioana Manolescu participated to a televised debate on fake news and the media in "Le Grand Barouf Numérique", a technology/society meet up organised by the city of Lille, in March 2018. The debate was broadcast by France 3 (https://france3-regions.francetvinfo.fr/hauts-de-france/nord-0/lille/debat-fake-news-medias-sont-ils-malades-1446365.html).
- Ioana Manolescu has participated to a debate on regulating the news media in "Les Journées de l'Economie", a national event organized in Lyon, in November 2018 (http://www.touteconomie.org/index.php?arc=bv1&manif=594)

9.3.3. Creation of media or tools for science outreach

Our research on computational fact-checking has been featured in a documentary (http://www.universcience.tv/video-tech-check-des-scientifiques-face-aux-fake-news-19501.html) by UniverScience.TV, the organization responsible of preparing scientific short movies next to Cité des Sciences. The movie has been featured in an article in Le Monde (https://www.lemonde.fr/sciences/video/2018/01/24/comment-la-science-aide-a-reperer-les-fake-news_5246356_1650684.html).

10. Bibliography

Major publications by the team in recent years

[1] R. BONAQUE, B. CAUTIS, F. GOASDOUÉ, I. MANOLESCU.*Social, Structured and Semantic Search*, in "International Conference on Extending Database Technology", Bordeaux, France, March 2016, https://hal.inria.fr/hal-01277939

[2] D. BURSZTYN, F. GOASDOUÉ, I. MANOLESCU.*Teaching an RDBMS about ontological constraints*, in "Very Large Data Bases", New Delhi, India, September 2016, https://hal.inria.fr/hal-01354592

[3] S. CAZALENS, P. LAMARRE, J. LEBLAY, I. MANOLESCU, X. TANNIER.*A Content Management Perspective on Fact-Checking*, in "The Web Conference 2018 - alternate paper tracks "Journalism, Misinformation and Fact Checking"", Lyon, France, April 2018, p. 565-574, https://hal.archives-ouvertes.fr/hal-01722666

[4] S. CEBIRIC, F. GOASDOUÉ, H. KONDYLAKIS, D. KOTZINOS, I. MANOLESCU, G. TROULLINOU, M. ZNEIKA.*Summarizing Semantic Graphs: A Survey*, in "The VLDB Journal", 2018, https://hal.inria.fr/hal-01925496

[5] E. HUANG, L. PENG, L. D. PALMA, A. ABDELKAFI, A. LIU, Y. DIAO.*Optimization for active learning-based interactive database exploration*, in "Proceedings of the VLDB Endowment (PVLDB)", September 2018, vol. 12, n^{o} 1, p. 71-84 [*DOI :* 10.14778/3275536.3275542], https://hal.inria.fr/hal-01969886

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

Publications of the year

Articles in International Peer-Reviewed Journal

[7] S. CAZALENS, J. LEBLAY, P. LAMARRE, I. MANOLESCU, X. TANNIER.*Computational Fact Checking: A Content Management Perspective*, in "Proceedings of the VLDB Endowment (PVLDB)", August 2018, vol. 11, n^{o} 12, p. 2110-2113 [*DOI :* 10.14778/3229863.3229880], https://hal.inria.fr/hal-01853067

[8] S. CEBIRIC, F. GOASDOUÉ, H. KONDYLAKIS, D. KOTZINOS, I. MANOLESCU, G. TROULLINOU, M. ZNEIKA.*Summarizing Semantic Graphs: A Survey*, in "The VLDB Journal", 2018, https://hal.inria.fr/hal-01925496

[9] C. CHANIAL, R. DZIRI, H. GALHARDAS, J. LEBLAY, M.-H. LE NGUYEN, I. MANOLESCU.*ConnectionLens: Finding Connections Across Heterogeneous Data Sources*, in "Proceedings of the VLDB Endowment (PVLDB)", 2018, vol. 11, 4 [*DOI :* 10.14778/3229863.3236252], https://hal.inria.fr/hal-01841009

[10] E. HUANG, L. PENG, L. D. PALMA, A. ABDELKAFI, A. LIU, Y. DIAO.*Optimization for active learning-based interactive database exploration*, in "Proceedings of the VLDB Endowment (PVLDB)", September 2018, vol. 12, n^{o} 1, p. 71-84 [*DOI :* 10.14778/3275536.3275542], https://hal.inria.fr/hal-01969886

International Conferences with Proceedings

[11] M. Buron, F. Goasdoué, I. Manolescu, M.-L. Mugnier.*Rewriting-Based Query Answering for Semantic Data Integration Systems*, in "34ème Conférence sur la Gestion de Données – Principes, Technologies et Applications (BDA 2018)", Bucarest, Romania, October 2018, https://hal.archives-ouvertes.fr/hal-01927282

[12] T.-D. Cao, I. Manolescu, X. Tannier.*Extracting Linked Data from statistic spreadsheets*, in "Conférence sur la Gestion de Données – Principes, Technologies et Applications", Bucarest, Romania, October 2018, https://hal.inria.fr/hal-01915148

[13] T.-D. Cao, I. Manolescu, X. Tannier.*Searching for Truth in a Database of Statistics*, in "WebDB 2018 - 21st International Workshop on the Web and Databases", Houston, United States, June 2018, p. 1-6, https://hal.inria.fr/hal-01745768

[14] C. Chanial, R. Dziri, H. Galhardas, J. Leblay, M.-H. Le Nguyen, I. Manolescu.*ConnectionLens: Finding Connections Across Heterogeneous Data Sources*, in "34ème Conférence sur la Gestion de Données – Principes, Technologies et Applications", Bucarest, Romania, October 2018, https://hal.inria.fr/hal-01968418

[15] P. Guzewicz, I. Manolescu.*Quotient RDF Summaries Based on Type Hierarchies*, in "DESWeb'2018 - Data Engineering meets the Semantic Web 2018", Paris, France, April 2018, https://hal.inria.fr/hal-01721163

[16] J. Leblay, I. Manolescu, X. Tannier.*Computational fact-checking: Problems, state of the art, and perspectives*, in "The Web Conference", Lyon, France, The Web Conference 2018, April 2018, https://hal.inria.fr/hal-01791232

[17] E. Pietriga, H. Gözükan, C. Appert, M. Destandau, Š. Čebirić, F. Goasdoué, I. Manolescu.*Browsing Linked Data Catalogs with LODAtlas*, in "ISWC 2018 - 17th International Semantic Web Conference", Monterey, United States, Springer, October 2018, p. 137-153, https://hal.inria.fr/hal-01827766

Conferences without Proceedings

[18] S. Cazalens, P. Lamarre, J. Leblay, I. Manolescu, X. Tannier.*A Content Management Perspective on Fact-Checking*, in "The Web Conference 2018 - alternate paper tracks "Journalism, Misinformation and Fact Checking"", Lyon, France, April 2018, p. 565-574, https://hal.archives-ouvertes.fr/hal-01722666

[19] F. Song, B. Zhou, Q. Sun, W. Sun, S. Xia, Y. Diao.*Anomaly Detection and Explanation Discovery on Event Streams*, in "BIRTE2018", RIO, Brazil, August 2018, https://hal.inria.fr/hal-01970660

Research Reports

[20] E. Huang, L. Peng, L. Di Palma, A. Abdelkafi, A. Liu, Y. Diao.*Optimization for Active Learning-based Interactive Database Exploration*, Ecole Polytechnique ; University of Massachusetts Amherst, 2018, https://hal.inria.fr/hal-01870560

[21] Š. Čebirić, F. Goasdoué, P. Guzewicz, I. Manolescu.*Compact Summaries of Rich Heterogeneous Graphs*, Inria Saclay ; Université Rennes 1, July 2018, n° RR-8920, p. 1-40, https://hal.inria.fr/hal-01325900

Other Publications

[22] I. CZERESNIA ETINGER.*Summary-based optimization in semantic graph databases*, March 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01742495

[23] P. GUZEWICZ.*Internship report: Quotient RDF graph summarization*, Inria Saclay ; Telecom ParisTech, September 2018, https://hal.inria.fr/hal-01879898

[24] I. MANOLESCU.*Democracy Big Bang: What data management can(not) do for journalism*, October 2018, 34ème Conférence sur la Gestion de Données – Principes, Technologies et Applications, https://hal.inria.fr/hal-01968347

[25] I. MANOLESCU.*Structural Summarization of Semantic Graphs*, June 2018, Extended Semantic Web Conference (ESWC) , https://hal.inria.fr/hal-01808737

Project-Team COMETE

Concurrency, Mobility and Transactions

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

IN PARTNERSHIP WITH:
CNRS

Ecole Polytechnique

RESEARCH CENTER
Saclay - Île-de-France

THEME
Security and Confidentiality

Table of contents

Project-Team COMETE

Creation of the Project-Team: 2008 January 01

Keywords:

Computer Science and Digital Science:

A2.1.1. - Semantics of programming languages
A2.1.5. - Constraint programming
A2.1.6. - Concurrent programming
A2.1.9. - Synchronous languages
A2.4.1. - Analysis
A2.4.2. - Model-checking
A3.4. - Machine learning and statistics
A4.1. - Threat analysis
A4.5. - Formal methods for security
A4.8. - Privacy-enhancing technologies

Other Research Topics and Application Domains:

B6.1. - Software industry
B6.6. - Embedded systems
B9.5.1. - Computer science
B9.10. - Privacy

1. Team, Visitors, External Collaborators

Research Scientists

Catuscia Palamidessi [Team leader, Inria, Senior Researcher]
Frank Valencia [CNRS, Researcher]
Konstantinos Chatzikokolakis [CNRS, Researcher, external member since Sep 2018 (détachement at the Univ. of Athens)]

Technical Staff

Ehab Elsalamouny [Inria, from Nov 2018, STIC]

PhD Students

Natasha Fernandes [Macquarie University, from Jul 2018]
Anna Pazii [Ecole Polytechnique]
Tymofii Prokopenko [Inria, until June 2018, Digicosme]
Santiago Quintero [Ecole Polytechnique, from Oct 2018]
Marco Romanelli [Inria, CORDI-S]

Post-Doctoral Fellows

Valentina Castiglioni [Inria]
Ali Kassem [Inria, until Aug 2018]

Visiting Scientists

Mario Ferreira Alvim Junior [Federal University of Minas Gerais, Brazil, Dec 2018]
Yusuke Kawamoto [AIST, Japan, Mar 2018 and Nov-Dec 2018]
Sergio Ramirez [Universidad Javeriana Cali, Colombia, from May 2018 until Jun 2018]
Carlos Olarte [Universidade Federal do Rio Grande do Norte, from Oct 2018 until Dec 2018]

Camilo Rueda [Universidad Javeriana Cali, Colombia, May 2018 and Nov 2018]

Administrative Assistants

Jessica Gameiro [Inria, until Apr 2018]

Maria Agustina Ronco [Inria, from May 2018]

2. Overall Objectives

2.1. Overall Objectives

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

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

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

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

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

3. Research Program

3.1. Probability and information theory

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

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

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

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

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

3.2. Expressiveness of Concurrent Formalisms

Participants: Catuscia Palamidessi, Frank Valencia.

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

3.3. Concurrent constraint programming

Participants: Frank Valencia, Santiago Quintero.

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

Our research in ccp develops along the following two lines:

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

3.4. Model checking

Participants: Konstantinos Chatzikokolakis, Catuscia Palamidessi.

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

4. Application Domains

4.1. Security and privacy

Participants: Catuscia Palamidessi, Konstantinos Chatzikokolakis, Ehab Elsalamouny, Ali Kassem, Anna Pazii, Marco Romanelli, Natasha Fernandes.

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

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

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

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

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

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

5. New Software and Platforms

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

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

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

The library's development continued in 2018 with several new added features. 82 new commits were pushed to the project's git repository during this year. The new functionality was directly applied to the exeperimental results of several publications of the team (QEST'18, Entropy'18, POST'18, CSF'18).

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

5.2. F-BLEAU

KEYWORDS: Information leakage - Machine learning - Privacy

FUNCTIONAL DESCRIPTION: F-BLEAU is a tool for estimating the leakage of a system about its secrets in a black-box manner (i.e., by only looking at examples of secret inputs and respective outputs). It considers a generic system as a black-box, taking secret inputs and returning outputs accordingly, and it measures how much the outputs "leak" about the inputs.

F-BLEAU is based on the equivalence between estimating the error of a Machine Learning model of a specific class and the estimation of information leakage.

This code was also used for the experiments of a paper under submission, on the following evaluations: Gowalla, e-passport, and side channel attack to finite field exponentiation.

RELEASE FUNCTIONAL DESCRIPTION: First F-BLEAU release. Supports frequentist and k-NN estimates with several parameters, and it allows stopping according to delta-convergence criteria.

- Contact: Konstantinos Chatzikokolakis
- URL: https://github.com/gchers/fbleau

5.3. Location Guard

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

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

Location Guard is by now a stable tool with a large user base. No new features were added in 2018, however the tool is still actively maintained, and several issues have beed fixed during this year (new geocoder API, manual installation method for Opera users, etc).

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

5.4. dspacenet

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

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

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

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

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

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

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

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

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

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

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

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

6. New Results

6.1. Foundations of information hiding

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

6.1.1. Secure Information Flow and Game Theory

In the inference attacks studied in Quantitative Information Flow (QIF), the attacker typically tries to interfere with the system in the attempt to increase its leakage of secret information. The defender, on the other hand, typically tries to decrease leakage by introducing some controlled noise. This noise introduction can be modeled as a type of protocol composition, i.e., a probabilistic choice among different protocols, and its effect on the amount of leakage depends heavily on whether or not this choice is visible to the attacker. In [21], [11], we considered operators for modeling visible and hidden choice in protocol composition, and we studied their algebraic properties. We then formalized the interplay between defender and attacker in a game-theoretic framework adapted to the specific issues of QIF, where the payoff is information leakage. We considered various kinds of leakage games, depending on whether players act simultaneously or sequentially, and on whether or not the choices of the defender are visible to the attacker. In the case of sequential games, the choice of the second player is generally a function of the choice of the first player, and his/her probabilistic choice can be either over the possible functions (mixed strategy) or it can be on the result of the function (behavioral strategy). We showed that when the attacker moves first in a sequential game with a hidden choice, then behavioral strategies are more advantageous for the defender than mixed strategies. This contrasts with the standard game theory, where the two types of strategies are equivalent. Finally, we established a hierarchy of these games in terms of their information leakage and provide methods for finding optimal strategies (at the points of equilibrium) for both attacker and defender in the various cases.

6.1.2. The additive capacity problem for Quantitative Information Flow

Preventing information leakage is a fundamental goal in achieving confidentiality. In many practical scenarios, however, eliminating such leaks is impossible. It becomes then desirable to quantify the severity of such leaks and establish bounds on the threat they impose. Aiming at developing measures that are robust wrt a variety of operational conditions, a theory of channel capacity for the g-leakage model was developed in [25], providing solutions for several scenarios in both the multiplicative and the additive setting. In [16] we continued this line of work by providing substantial improvements over the results of [25] for additive leakage. The main idea of employing the Kantorovich distance remains, but it is now applied to quasimetrics, and in particular the novel “convex-separation” quasimetric. The benefits were threefold: first, it allowed to maximize leakage over a larger class of gain functions, most notably including the one of Shannon. Second, a solution was obtained to the problem of maximizing leakage over both priors and gain functions, left open in [25]. Third, it allowed to establish an additive variant of the " Miracle " theorem from [26].

6.1.3. Local Differential Privacy and Statistical Utility

Local differential privacy (LDP) is a variant of differential privacy (DP) where the noise is added directly on the individual records, before being collected. The main advantage with respect to DP is that we do not need a trusted third party to collect and sanitise the sensitive data of the user. The main disadvantage is that the trade-off between privacy and utility is usually worse than in DP, and typically to retrieve reasonably good statistics from the locally sanitised data it is necessary to have access to a huge collection of them. In [22], we focused on the problem of estimating the counting queries on numerical data, and we proposed a variant of LDP based on the addition of geometric noise. Such noise function is known to have appealing properties in the case of counting queries. In particular, it is universally optimal for DP, i.e., it provides the best utility for a given level of DP, regardless of the side knowledge of the attacker. We explored the properties of geometric noise for counting queries in the LDP setting, and we conjectured an optimality property, similar to the one that holds in the DP setting. In [15] we proposed a variant of LDP suitable for metric spaces, such as location data or energy consumption data, and we showed that it provides a better utility, for the same level of privacy, then the other known LPD mechanisms.

6.1.4. Information-Theoretic Methods for Feature Selection in Machine Learning

The identification of the “best” features for classification is a problem of increasing importance in machine learning. The size of available datasets is becoming larger and larger, both in terms of samples and in terms of features of the samples, and keeping the dimensionality of the data under control is necessary for avoiding an explosion of the training complexity and for the accuracy of the classification. The known methods for reducing the dimensionality can be divided in two categories: those which transform the feature space by reshaping the original features into new ones (feature extraction), and those which select a subset of the features (feature selection). Several proposals for feature selection have successfully applied concepts and techniques from information theory. In [19] we proposed a new information-theoretic algorithm for ordering the features according to their relevance for classification. The novelty of our proposal consisted in adopting Rényi min-entropy instead of the commonly used Shannon entropy. In particular, we adopted a notion of conditional min-entropy that has been recently proposed in the field of security and privacy, and that avoids the anomalies of previously-attempted information-theoretic definitions. This notion is strictly related to the Bayes error, which is a promising property for achieving accuracy in the classification. We evaluated our method on various classifiers and datasets, and we showed that it compares favorably to the corresponding one based on Shannon entropy.

6.1.5. A Logical Characterization of Differential Privacy via Behavioral Metrics

Differential privacy (DP) is a formal definition of privacy ensuring that sensitive information relative to individuals cannot be inferred by querying a database. In [18], we exploited a modeling of this framework via labeled Markov Chains (LMCs) to provide a logical characterization of differential privacy: we considered a probabilistic variant of the Hennessy-Milner logic and we defined a syntactical distance on formulae in it measuring their syntactic disparities. Then, we defined a trace distance on LMCs in terms of the syntactic distance between the sets of formulae satisfied by them. We proved that such distance corresponds to the level of privacy of the LMCs. Moreover, we used the distance on formulae to define a real-valued semantics for them, from which we obtained a logical characterization of weak anonymity: the level of anonymity is measured in terms of the smallest formula distinguishing the considered LMCs. Then, we focused on bisimulation semantics on nondeterministic probabilistic processes and we provide a logical characterization of generalized bisimulation metrics, namely those defined via the generalized Kantorovich lifting. Our characterization is based on the notion of mimicking formula of a process and the syntactic distance on formulae, where the former captures the observable behavior of the corresponding process and allows us to characterize bisimilarity. We showed that the generalized bisimulation distance on processes is equal to the syntactic distance on their mimicking formulae. Moreover, we used the distance on mimicking formulae to obtain bounds on differential privacy.

6.1.6. Probability and Nondeterminism in Process Calculi from a Logical Perspective

Behavioral equivalences and modal logics have been successfully employed for the specification and verification of communicating concurrent systems, henceforth processes. The former ones, in particular the family of bisimulations, provide a simple and elegant tool for the comparison of the observable behavior of processes. The latter ones allow for an immediate expression of the desired properties of processes. Since the work on the Hennessy-Milner logic (HML), these two approaches are connected by means of logical characterizations of behavioral equivalences: two processes are behaviorally equivalent if and only if they satisfy the same formulae in the logic. Hence, the characterization of an equivalence subsumes both the fact that the logic is as expressive as the equivalence and the fact that the equivalence preserves the logical properties of processes. However, the connection between behavioral equivalences and modal logics goes even further: modal decomposition of formulae exploits the characterization of an equivalence to derive its compositional properties. Roughly speaking, the definition of the semantic behavior of processes by means of the Structural Operational Semantics (SOS) framework allowed for decomposing the satisfaction problem of a formula for a process into the verification of the satisfaction problem of certain formulae for its subprocesses. In [12] we extended the SOS-driven decomposition approach to processes in which the nondeterministic behavior coexists with probability. To deal with the probabilistic behavior of processes, and thus with the decomposition of formulae characterizing it, we introduced a SOS-like machinery allowing for the specification of the behavior of open distribution terms. By our decomposition, we obtained (pre)congruence formats for probabilistic bisimilarity, ready similarity and similarity.

The combination of nondeterminism and probability in concurrent systems leads to different interpretations of process behavior. If we restrict our attention to linear properties only, we can identify three main approaches to trace and testing semantics: the trace distributions, the trace-by-trace and the extremal probabilities approaches. In [17] we proposed novel notions of behavioral metrics that are based on the three classic approaches above, and that can be used to measure the disparities in the linear behavior of processes wrt. trace and testing semantics. We studied the properties of these metrics, like non-expansiveness, and we compare their expressive powers.

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. Real-time Rewriting Logic Semantics for Spatial Concurrent Constraint Programming

In [20] we used rewriting logic for specifying and analyzing a calculus for concurrent constraint programming (ccp) processes combining spatial and real-time behavior. These processes can run processes in different computational spaces (e.g., containers) while subject to real-time requirements (e.g., upper bounds in the execution time of a given operation), which can be specified with both discrete and dense linear time. The real-time rewriting logic semantics is fully executable in Maude with the help of rewriting modulo SMT: partial information (i.e., constraints) in the specification is represented by quantifier-free formulas on the shared variables of the system that are under the control of SMT decision procedures. The approach is used to symbolically analyze existential real-time reachability properties of process calculi in the presence of spatial hierarchies for sharing information and knowledge.

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

In [14] spatial constraint systems are used to give an abstract characterization of the notion of normality in modal logic and to derive right inverse/reverse operators for modal languages. In particular, a necessary and sufficient condition for the existence of right inverses is identified and the abstract notion of normality is shown to correspond to the preservation of finite suprema. Furthermore, a taxonomy of normal right inverses is provided, identifying the greatest normal right inverse as well as the complete family of minimal right inverses. These results were applied to existing modal languages such as the weakest normal modal logic, Hennessy-Milner logic, and linear-time temporal logic. Some implications of these results were also discussed in the context of modal concepts such as bisimilarity and inconsistency invariance.

6.2.3. Observational and Behavioural Equivalences for Soft Concurrent Constraint Programming

In [13] we presented a labelled semantics for Soft Concurrent Constraint Programming (SCCP), a meta-language where concurrent agents may synchronise on a shared store by either posting or checking the satisfaction of (soft) constraints. SCCP generalises the classical formalism by parametrising the constraint system over an order-enriched monoid, thus abstractly representing the store with an element of the monoid, and the standard unlabelled semantics just observes store updates. The novel operational rules were shown to offer a sound and complete co-inductive technique to prove the original equivalence over the unlabelled semantics. Based on this characterisation, we provided an axiomatisation for finite agents.

7. Partnerships and Cooperations

7.1. Regional Initiatives

7.1.1. OPTIMEC

Project title: Optimal Mechanisms for Privacy Protection

Funded by: DigiCosme

Duration: September 2016 - July 2018

Coordinator: Catuscia Palamidessi, Inria Saclay, EPI Comète

Other PI's: Serge Haddad, ENS Cachan.

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

7.1.2. SUPREME

Project title: Statistical-Utility Preserving Methods for Privacy Protection

Funded by: Département STIC

Duration: 2018 - 2019

Coordinator: Catuscia Palamidessi, Inria Saclay, EPI Comète

Other PI's: Serge Haddad, ENS Cachan.

Abstract: In this project we study the theoretical foundations, methods and tools to protect the privacy of the individuals under certain constraints. In particular we focus on mechanisms that: (1) are robust with respect to combination of information from different sources, (2) can be applied directly by the user, thus avoiding the need of a trusted party, and (3) provide an optimal trade-off between privacy and utility.

7.2. National Initiatives

7.2.1. REPAS

Program: ANR Blanc

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

Duration: October 2016 - September 2021

Coordinator: Catuscia Palamidessi, Inria Saclay, EPI Comète

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

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

7.2.2. MAGIC

Program: PEPS I3A

Project title: Machine Games for Information Protection

Duration: February 2018 - December 2018

Coordinator: Konstantinos Chatzikokolakis, CNRS (EPI Comète) and Ecole Polytechnique

Other PI's and partner institutions: Giovanni Cherubin, EPFL, Switzerland. Serge Haddad, ENS Cachan.

Abstract: In this project, we study a Machine Learning approach to develop methods for the Protection of Private Information. The idea is based on the Generative Adversarial Network (GAN) paradigm: the defender and the attacker are modeled as two adversaries in a game, where the payoff is the attacker's acquisition of the user's private data by exploiting the system vulnerabilities, side information, and probabilistic inference.

7.3. International Initiatives

7.3.1. Inria Associate Teams

7.3.1.1. LOGIS

Title: Logical and Formal Methods for Information Security

Inria principal investigator: Konstantinos Chatzikokolakis

International Partners:

- Mitsuhiro Okada, Keio University (Japan)
- Yusuke Kawamoto, AIST (Japan)
- Tachio Terauchi, JAIST (Japan)

Masami Hagiya, University of Tokyo (Japan)

Start year: January 2016 - December 2018

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

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

7.3.2. Participation in International Programs

7.3.2.1. CLASSIC

Program: Colciencias - Conv. 712.

Project acronym: CLASSIC.

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

Duration: Oct 2016 - Oct 2019.

URL: http://goo.gl/Gv6Lij

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

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

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

7.3.2.2. EPIC

Program: STIC-Amsud.

Project acronym: EPIC.

Project title: EPistemic Interactive Concurrency/

Duration: Oct 2016 - Oct 2018.

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

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

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

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

7.3.2.3. FACTS

Program: ECOS NORD.

Project acronym: FACTS.

Project title: Foundational Approach to Cognition in Today?s Society.

Duration: Jan 1 2019 - Dec 31, 2021.

URL: https://goo.gl/zVhg32

Coordinator: Frank Valencia, Ecole Polytechnique.

Other PI's and partner institutions: Jean-Gabriel Ganascia LIP6, Sorbonne University and Camilo Rueda, Universidad Javeriana de Cali, Colombia.

Abstract: This projects aims at studying the phenomenon of “Group Polarization”; the tendency for a group to learn or acquire beliefs or to make decisions that are more extreme than the initial inclinations of its members.

7.3.3. Inria International Partners

7.3.3.1. PriDat

Project title: Privacy-Friendly Data Analytics

Funded by: Siebel Energy Institute

Duration: September 2018 - August 2019

Coordinator: Catuscia Palamidessi, Inria Saclay, EPI Comète

Other PI’s: Giovanni Cherubin, EPFL, Switzerland. Moreno Falaschi, University of Siena, Italy. Mario Ferreira, Federal University of Minas Gerais, Brazil.

Abstract: The objective of this project is to develop methodologies for protecting the privacy of individuals while letting their data be collected and used for analytical purposes.

7.3.3.2. Informal International Partners

Geoffrey Smith, Florida International University, USA

Carroll Morgan, NICTA , Australia

Annabelle McIver, Maquarie University, Australia

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

Camilo Rueda, Professor, Universidad Javeriana de Cali, Colombia

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

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

7.4. International Research Visitors

7.4.1. Visits of International Scientists

Mario Ferreira Alvim Junior, Assistant Professor, Federal University of Minas Gerais, Brazil. Dec 2018

Borja de Balle Pigem. Sr. Machine Learning Scientist. Amazon, UK. Dec 2018

Takao Murakami, Assistant Professor, National Institute of Advanced Industrial Science and Technology (AIST), Japan. Dec 2018

Yusuke Kawamoto, Assistant Professor, National Institute of Advanced Industrial Science and Technology (AIST), Japan. March 2018 and Nov-Dec 2018

Carlos Olarte, Assistant Professor, Universidade Federal do Rio Grande do Norte, Brazil. Oct-Dec 2018

Daniele Gorla, Professor, University of Rome “La Sapienza”. Aug - Sep 2018.

Giovanna Broccia, PhD student, University of Pisa, Italy, June 2018

Camilo Rueda, Professor, Universidad Javeriana de Cali, Colombia. May 2018 and Nov 2018

Prakash Panangaden, University of McGill, Montreal, Canada. Feb 2018

7.4.2. Internships

Haoteng Yin. Academy for Advanced Interdisceplinary Studies, Peking University. From June 2018 until Sept 2018.

Kacem Kefki. University of Paris Saclay. From June 2018 until July 2018.

Arthur Américo. Universidade Federal de Minas Gerais. From April 2018 until June 2018.

Noémie Fong. ENS Paris. From April 2018 until Sept 2019.

Pedro Bahamondes. Ecole Polytechnique. From Sept 2017 until March 2018.

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

Jason Lopez, Universidad Javeriana de Cali, Colombia. From May until Agost 2018.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific events organisation

8.1.1.1. Member of the organizing committee

Catuscia Palamidessi is member of:

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

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

The Organizing Committee of LICS, the ACM/IEEE Symposium on Logic in Computer Science. 2014-18.

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

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

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

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

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

The IFIP Working Group 1.8 – Concurrency Theory. Since 2005.

Frank D. Valencia is member of:

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

Konstantinos Chatzikokolakis is member of:

The steering committee of the Privacy Enhancing Technologies Symposium. Since 2018.

8.1.2. Scientific events selection committee

8.1.2.1. Chair of conference program committee

Konstantinos Chatzikokolakis:

is serving as PC chair (with Carmela Troncoso as co-chair) of PETS 2019: The 19th Privacy Enhancing Technologies Symposium, July 16-20, 2019, Stockholm, Sweden.

8.1.2.2. Member of conference program committees

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

PETS 2019. The 19th Privacy Enhancing Technologies Symposium. Stockholm, Sweden, 16–20 July, 2019.

LICS 2019. The Thirty-Fourth Annual ACM/IEEE Symposium on Logic in Computer Science. Vancouver, Canada, 24–27 June 2019.

CSF 2019. The 32nd IEEE Computer Security Foundations Symposium. Hoboken, NJ, USA, June 24-27, 2019.

SAC 2019 (Security track). The 34th ACM/SIGAPP Symposium On Applied Computing. Limassol, Cyprus, 8-12 April 2019.

FACS 2018. The 15th International Conference on Formal Aspects of Component Software. Pohang, South Corea, 10-12 October 2018.

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

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

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

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

PPML 2018. Privacy Preserving Machine Learning (NeurIPS 2018 Workshop). Montréal, Canada, 8 December 2018.

APVP 2018. Atelier sur la Protection de la Vie Privée. Porquerolles, France, 3-6 juin 2018.

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

BMDA 2018: Workshop on Big Mobility Data Analytics

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

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

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

CP-ICLP-SAT-DP-18. Doctoral Program of the 23rd International Conference on Principles and Practice of Constraint Programming.

CONCUR 2019. The 30th International Conference on Concurrency Theory. Amsterdam, the Netherlands, August 26-31, 2019.

AAMAS 2019. International Conference on Autonomous Agents and Multiagent Systems. Montreal, Canada, 13th-17th of May 2019.

8.1.3. Journals

8.1.3.1. Member of the editorial board

Catuscia Palamidessi is:

Member of the Editorial Board of the Proceedings on Privacy Enhancing Technologies (PoPETs), published by De Gruyter. Since 2017.

Member of the Editorial Board of Mathematical Structures in Computer Science, published by the Cambridge University Press. Since 2006.

Member of the Editorial Board of Acta Informatica, published by Springer. Since 2015.

Member of the Editorial Board of the Electronic Notes of Theoretical Computer Science, published by Elsevier Science. Since 2000.

Member of the Editorial Board of LIPIcs: Leibniz International Proceedings in Informatics, Schloss Dagstuhl–Leibniz Center for Informatics. Since 2014.

Konstantinos Chatzikokolakis is:

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

8.1.3.2. Reviewing

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

8.1.4. Other Editorial Activities

Catuscia Palamidessi is/has been:

Co-editor (with Anca Muscholl and Anuj Dawar) of the special issue of Logical Methods in Computer Science dedicated to selected papers of ICALP 2017.

Co-editor (with Alexandra Silva and Natarajan Shankar) of the special issue of Logical Methods in Computer Science dedicated to selected papers of LICS 2015 and LICS 2016.

Frank D. Valencia has been:

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

8.1.5. Participation in other committees

Catuscia Palamidessi has been serving in the following committees:

Member of the panel for the Research Evaluation for Development 2019 (RED19) of the Department of Computer Science and Engineering at the University of Gothenburg, Sweden.

Chair of the Nominating Committee for the 2019 renewal of the office holders of SIGLOG, the ACM Special Interest Group on Logic and Computation.

Member of the evaluation panel for the SU-ICT-03-2018: “Dynamic countering of cyber-attacks” - H2020 Work Programme 2018-2020.

Member of the evaluation panel for the program IKTPLUSS on Digital Security, Research Council of Norway (2018).

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

Reviewer for the projects proposal for the program PRIN, sponsored by the Italian MIUR (“Ministero dell’Istruzione, dell’Università e della Ricerca”). Since 2005.

President of the selection committee for the EATCS Best Paper Award. From 2006 until 2018.

Member of the EAPLS PhD Award Committee. From 2010.

8.1.6. Invited talks

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

CSL 2018. The 27th Computer Science Logic Annual Conference. Birmingham, UK, 4–7 September 2018.

CSF 2018. The 31st IEEE Computer Security Foundations Symposium, Oxford, UK, 9-12 July 2018.

PROOFS 2018 (Keynote speaker). The 7th International Workshop on Security Proofs for Embedded Systems. Amsterdam, The Netherlands, 13 September 2018.

PiMLAI 2018 Privacy in Machine Learning and Artificial Intelligence (FAIM 2018 Workshop). Stockholm, Sweden, 15 July 2018.

Bernoulli Symposium. Opening Symposium of the new institute for Artificial Intelligence, Mathematics, and Computer Science of the University of Groningen. Groningen, The Netherlands, 1 November 2018.

Konstantinos Chatzikokolakis has given invited talks at the following conference:

QEST 2018. 15th International Conference on Quantitative Evaluation of SysTems, Beijing, China, September 4-7, 2018.

8.1.7. Service

Catuscia Palamidessi has served as:

Member of the committee for the assignment of the Inria International Chairs. From 2017.

Frank Valencia has served as:

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

Konstantinos Chatzikokolakis has served as:

Member of the hiring committee for the "poste enseignants-chercheur Gaspard Monge", Ecole Polytechnique, 2018.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Master : Frank D. Valencia has been teaching the undergraduate course "Computability", 45 hours, at the Pontificia Universidad Javeriana de Cali, Colombia. July 27 - Nov 1, 2018.

Master : Frank D. Valencia has been teaching the masters course "Foundations of Computer Science", 45 hours, at the Pontificia Universidad Javeriana de Cali, Colombia. Jan 27 - Jun 1, 2018

Master : Konstantinos Chatzikokolakis has been teaching the masters course "Privacy Technologies", 40 hours, at the University of Athens, Greece. Oct - Dec, 2018.

8.2.2. Supervision

PhD in progress (2018-) Natasha Fernandez. Co-supervised Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Annabelle McIver. Thesis subject: Privacy Protection Methods for Textual Documents.

PhD in progress (2018-) Santiago Quintero. Co-supervised by Frank Valencia and Catuscia Palamidessi. Thesis Subject: Foundations of Group Polarization.

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

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

PhD terminated (2016-18) Tymofii Prokopenko. Ecole Polytechnique and ENS Cachan. Grant Digiteo-Digicosme. Co-supervised by Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Serge Haddad (ENS Cachan). The PhD was terminated due to the lack of progress.

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

8.2.3. Juries

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

Vittoria Nardone (University of Sannio, Italy). PhD thesis reviewer. Title of the thesis: *Formal Methods for Android Applications*. Supervised by Antonella Santone. Defended in January 2019.

Antoine Dallon (ENS Paris-Saclay). PhD thesis reviewer and member of the committee board at the PhD defense. Title of the thesis: *Verification of indistinguishability properties in cryptographic protocols - Small attacks and efficient decision with SAT-Equiv*. Supervised by Veronique Cortier and Stephanie Delaune. Defended on November 26, 2018.

David Mestel (University of Oxford). PhD thesis reviewer and member of the committee board at the PhD defense. Title of the thesis: *Quantifying information flow*. Supervised by Bill Roscoe. Defended on October 26, 2018.

Jun Wang (University of Luxembourg). PhD thesis reviewer and member of the committee board at the PhD defense. Title of the thesis: *Privacy-preserving recommender systems facilitated by machine learning approach*. Supervised by Qiang Tang and Peter Ryan. Defended on October 19, 2018.

Hamid Ebadi (Chalmers University, Sweden). Member of the committee board at the PhD defense. Title of the thesis: *Dynamic Enforcement of Differential Privacy*. Supervised by David Sands. Defended on March 5, 2018.

Catuscia Palamidessi has been examiner of the following habilitation thesis:

Elham Kashefi (LIP6, CNRS, France). Title of the thesis: *Verification of Quantum Computing*. Defended on February 8, 2018.

8.2.4. Other didactical duties

Catuscia Palamidessi has been:

Member of the advising committee for Hamid Ebadi, PhD student supervised by David Sands, Chalmers University, Sweden. From 2014 until 2018.

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

Member of the advising committee for the PhD of Jun Wang (PhD student supervised by Qiang Tang and Peter Ryan), University of Luxembourg. From 2014 until 2018.

8.3. Popularization

8.3.1. Education

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

Catuscia Palamidessi has been:

- Invited speaker at PLMW@POPL 2019, the Programming Logic Mentoring Workshop 2019 (affiliated to POPL 2019). This workshop aims at encouraging graduate students and senior undergraduate students to pursue careers in programming language research, and at educating them on the research career.
- A participant in the round table at the FLOC Women in Logic workshop, a workshop organized to encourage women's presence in the logic community. Oxford, UK, 8 July 2018.

8.3.2. Interventions

Catuscia Palamidessi and Frank Valencia have supervised a group of high school children in stage d'observation. April 2018.

Catuscia Palamidessi has given an invited talk at:

- JNIM 2018. Journées Nationales 2018 du GDR Informatique Mathématique (Journée du 6 Avril en Hommage à Maurice Nivat). Palaiseau, France, 3-6 April 2018.

9. Bibliography

Major publications by the team in recent years

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

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

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

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

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

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

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

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

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

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

Publications of the year

Articles in International Peer-Reviewed Journal

[11] M. S. ALVIM, K. CHATZIKOKOLAKIS, Y. KAWAMOTO, C. PALAMIDESSI.*A Game-Theoretic Approach to Information-Flow Control via Protocol Composition*, in "Entropy", May 2018, vol. 20, n^{o} 5, 382 [*DOI :* 10.3390/E20050382], https://hal.archives-ouvertes.fr/hal-01966862

[12] V. CASTIGLIONI, D. GEBLER, S. TINI.*SOS-based Modal Decomposition on Nondeterministic Probabilistic Processes*, in "Logical Methods in Computer Science", August 2018, vol. 14, n^{o} 2 [*DOI :* 10.23638/LMCS-14(2:18)2018], https://hal.inria.fr/hal-01966954

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

[14] M. GUZMÁN, S. PERCHY, C. RUEDA, F. VALENCIA.*Characterizing Right Inverses for Spatial Constraint Systems with Applications to Modal Logic*, in "Theoretical Computer Science", October 2018, vol. 744, n^{o} 56–77, https://hal.inria.fr/hal-01675010

Invited Conferences

[15] M. S. ALVIM, K. CHATZIKOKOLAKIS, C. PALAMIDESSI, A. PAZII.*Metric-based local differential privacy for statistical applications*, in "31st Computer Security Foundations Symposium (CSF 2018)", Oxford, United Kingdom, IEEE Computer Society, July 2018, p. 262-267, https://hal.archives-ouvertes.fr/hal-01966869

[16] K. CHATZIKOKOLAKIS.*On the Additive Capacity Problem for Quantitative Information Flow*, in "15th International Conference on Quantitative Evaluation of SysTems (QEST 2018)", Beijing, China, Lecture Notes in Computer Science, Springer, September 2018, vol. 11024, p. 1-19, https://hal.inria.fr/hal-01845330

International Conferences with Proceedings

[17] V. CASTIGLIONI.*Trace and Testing Metrics on Nondeterministic Probabilistic Processes*, in "Proceedings Combined 25th International Workshop on Expressiveness in Concurrency and 15th Workshop on Structural Operational Semantics and 15th Workshop on Structural Operational Semantics, (EXPRESS/SOS) 2018", Beijing, China, September 2018, vol. 276, p. 19-36 [*DOI :* 10.4204/EPTCS.276.4], https://hal.inria.fr/hal-01966950

[18] V. CASTIGLIONI, K. CHATZIKOKOLAKIS, C. PALAMIDESSI.*A Logical Characterization of Differential Privacy via Behavioral Metrics*, in "Formal Aspects of Component Software (FACS 2018)", Pohang, South Korea, K. BAE, P. C. ÖLVECZKY (editors), Springer, October 2018, vol. 11222, p. 75-96 [*DOI :* 10.1007/978-3-030-02146-7_4], https://hal.archives-ouvertes.fr/hal-01966870

[19] C. PALAMIDESSI, M. ROMANELLI.*Feature selection with Rényi min-entropy*, in "Artificial Neural Networks in Pattern Recognition - 8th IAPR TC3 Workshop (ANNPR 2018)", Siena, Italy, L. PANCIONI, F. SCHWENKER, E. TRENTIN (editors), Springer, September 2018, vol. 11081, p. 226-239, https://hal.archives-ouvertes.fr/hal-01830177

[20] S. RAMÍREZ, M. ROMERO, C. ROCHA, F. D. VALENCIA.*Real-time Rewriting Logic Semantics for Spatial Concurrent Constraint Programming*, in "Rewriting Logic and Its Applications - 12th International Workshop", Thessaloniki, Greece, V. RUSU (editor), Rewriting Logic and Its Applications, Springer, June 2018, vol. 11152, p. 226–244, https://hal.archives-ouvertes.fr/hal-01934953

Conferences without Proceedings

[21] M. S. ALVIM, K. CHATZIKOKOLAKIS, Y. KAWAMOTO, C. PALAMIDESSI.*Leakage and Protocol Composition in a Game-Theoretic Perspective*, in "7th International Conference on Principles of Security and Trust (POST 2018)", Thessaloniki, Greece, L. BAUER, R. KÜSTERS (editors), Springer, April 2018, vol. 10804, p. 134-159, https://hal.archives-ouvertes.fr/hal-01966871

[22] L. KACEM, C. PALAMIDESSI.*Geometric Noise for Locally Private Counting Queries*, in "Proceedings of the 13th Workshop on Programming Languages and Analysis for Security (PLAS 2018)", Toronto, Canada, ACM, 2018, p. 13-16 [*DOI :* 10.1145/3264820.3264827], https://hal.archives-ouvertes.fr/hal-01966872

Scientific Books (or Scientific Book chapters)

[23] M. S. ALVIM, K. CHATZIKOKOLAKIS, A. MCIVER, C. MORGAN, C. PALAMIDESSI, G. SMITH.*The Science of Quantitative Information Flow*, Springer, 2019, https://hal.inria.fr/hal-01971490

[24] M. LEUCKER, J. A. PÉREZ, C. RUEDA, F. D. VALENCIA.*Special Issue: Best Papers Presented at ICTAC 2015*, Mathematical Structures in Computer Science, Cambridge University Press, May 2018, https://hal.archives-ouvertes.fr/hal-01837891

References in notes

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

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

Project-Team COMMANDS

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

IN COLLABORATION WITH: Centre de Mathématiques Appliquées (CMAP)

RESEARCH CENTER
Saclay - Île-de-France

THEME
Optimization and control of dynamic systems

Table of contents

Project-Team COMMANDS

Creation of the Project-Team: 2009 January 01

Keywords:

Computer Science and Digital Science:

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

Other Research Topics and Application Domains:

B4.4. - Energy delivery
B4.4.1. - Smart grids
B7.1.2. - Road traffic
B7.1.3. - Air traffic
B7.2.1. - Smart vehicles

1. Team, Visitors, External Collaborators

Research Scientists

Joseph Frederic Bonnans [Team leader, Inria, Senior Researcher, HDR]
Pierre Martinon [Inria, Researcher]

Technical Staff

Liu Jinyan [Inria, until Jan 2018]

PhD Students

Guillaume Bonnet [Univ. Paris Saclay, from Oct 2018]
Saeed Hadikhanloo [Inria, until Jan 2018]
Pierre Lavigne [École Nationale Supérieure de Techniques Avancées, from Oct 2018]
Arthur Le Rhun [Ifpen, from Sep 2016]
Cédric Rommel [Safety Line, until Oct 2018]

Post-Doctoral Fellows

Justina Gianatti [Inria, from May 2018]
Saeed Hadikhanloo [Inria, from Feb 2018]

Visiting Scientist

Axel Kröner [Univ. Humboldt]

Administrative Assistants

Hanadi Dib [Inria, until Oct 2018]
Ines Dumontier [Inria, from Nov 2018]
Hélèna Kutniak [Inria, from Oct 2018]

2. Overall Objectives

2.1. Scientific directions

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

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

2.2. Industrial impact

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

- Space vehicle trajectories, in collaboration with CNES, the French space agency.
- Aeronautics, in collaboration with the startup Safety Line.
- Production, management, storage and trading of energy resources, in collaboration with Edf, ex-Gdf and Total.
- Energy management for hybrid vehicles, in collaboration with Renault and Ifpen.

We give more details in the Bilateral contracts section.

3. Research Program

3.1. Historical aspects

The roots of deterministic optimal control are the "classical" theory of the calculus of variations, illustrated by the work of Newton, Bernoulli, Euler, and Lagrange (whose famous multipliers were introduced in [28]), with improvements due to the "Chicago school", Bliss [20] during the first part of the 20th century, and by the notion of relaxed problem and generalized solution (Young [33]).

Trajectory optimization really started with the spectacular achievement done by Pontryagin's group [32] during the fifties, by stating, for general optimal control problems, nonlocal optimality conditions generalizing those of Weierstrass. This motivated the application to many industrial problems (see the classical books by Bryson and Ho [24], Leitmann [30], Lee and Markus [29], Ioffe and Tihomirov [27]).

Dynamic programming was introduced and systematically studied by R. Bellman during the fifties. The HJB equation, whose solution is the value function of the (parameterized) optimal control problem, is a variant of the classical Hamilton-Jacobi equation of mechanics for the case of dynamics parameterized by a control variable. It may be viewed as a differential form of the dynamic programming principle. This nonlinear first-order PDE appears to be well-posed in the framework of *viscosity solutions* introduced by Crandall and Lions [25]. The theoretical contributions in this direction did not cease growing, see the books by Barles [18] and Bardi and Capuzzo-Dolcetta [17].

3.2. Trajectory optimization

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

In the *full discretization approach*, general Runge-Kutta schemes with different values of control for each inner step are used. This allows to obtain and control high orders of precision, see Hager [26], Bonnans [21]. In the *indirect* approach, the control is eliminated thanks to Pontryagin's maximum principle. One has then to solve the two-points boundary value problem (with differential variables state and costate) by a single or multiple shooting method. The questions are here the choice of a discretization scheme for the integration of the boundary value problem, of a (possibly globalized) Newton type algorithm for solving the resulting finite dimensional problem in IR^n (n is the number of state variables), and a methodology for finding an initial point.

3.3. Hamilton-Jacobi-Bellman approach

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

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

For solving stochastic control problems, we studied the so-called Generalized Finite Differences (GFD), that allow to choose at any node, the stencil approximating the diffusion matrix up to a certain threshold [23]. Determining the stencil and the associated coefficients boils down to a quadratic program to be solved at each point of the grid, and for each control. This is definitely expensive, with the exception of special structures where the coefficients can be computed at low cost. For two dimensional systems, we designed a (very) fast algorithm for computing the coefficients of the GFD scheme, based on the Stern-Brocot tree [22].

4. Application Domains

4.1. Fuel saving by optimizing airplanes trajectories

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

4.2. Hybrid vehicles

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

4.3. Biological systems

2018 was the last year of IPL Algae in Silico in which we tackled the optimization of photobioreactors in turbid conditions. Our participation to IPL Cosy was strenghtened through the co-supervision of the PhD of E. Weill with team Lifeware (Inria and Pasteur), starting from september 2018, and focused on the oprimization of heterogenous populations of micro-organisms.

5. Highlights of the Year

5.1. Highlights of the Year

Variational analysis for options with stochastic volatility and multiple factors

Publication of the paper [3] in the SIAM J. finance. This paper clarifies the issue of well-posedness of some PDEs arising in finance.

A stochastic data-based traffic model applied to vehicles energy consumption estimation

Publication [10] of a new method for the analysis of road traffic, in relation with energy consumption.

6. New Software and Platforms

6.1. BOCOP

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

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

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

6.2. Bocop HJB

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

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

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

6.3. Bocop Avion

KEYWORDS: Optimization - Aeronautics
FUNCTIONAL DESCRIPTION: Optimize the climb speeds and associated fuel consumption for the flight planning of civil airplanes.
NEWS OF THE YEAR: Improved atmosphere model 2D interpolations for temperature and wind data

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

6.4. Bocop HJB Avion

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

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

7. New Results

7.1. Optimal control of ODEs

7.1.1. Optimal health insurance design

In [7] we analyze the design of optimal medical insurance under ex post moral hazard, i.e., when illness severity cannot be observed by insurers and policyholders decide for themselves on their health expenditures. The trade-off between ex anterisk sharing and ex post incentive compatibility is analyzed in an optimal revelation mechanism under hidden information and risk aversion. The optimal contract provides partial insurance at the margin, with a deductible when insurers' rates are affected by a positive loading, and it may also include an upper limit on coverage. The potential to audit the health state leads to an upper limit on out-of-pocket expenses.

Health insurance with audit. The 'out-of-pocket' expense $(m - I)$ remains bounded.

7.1.2. Optimal Battery Aging: an Adaptive Weights Dynamic Programming Algorithm

In [5] we present an algorithm to handle the optimization over a long horizon of an electric microgrid including a battery energy storage system. While the battery is an important and costly component of the microgrid, its aging process is often not taken into account by the Energy Management System, mostly because of modeling and computing challenges. We address the computing aspect by a new approach combining dynamic programming, decomposition and relaxation techniques. We illustrate this 'adaptive weight' method with numerical simulations for a toy microgrid model. Compared to a straightforward resolution by dynamic programming, our algorithm decreases the computing time by more than one order of magnitude, can be parallelized, and allows for online implementations. We believe that this approach can be used for other applications presenting fast and slow variables.

Optimal battery aging. Comparison of brute-force and adaptive weights algorithm.

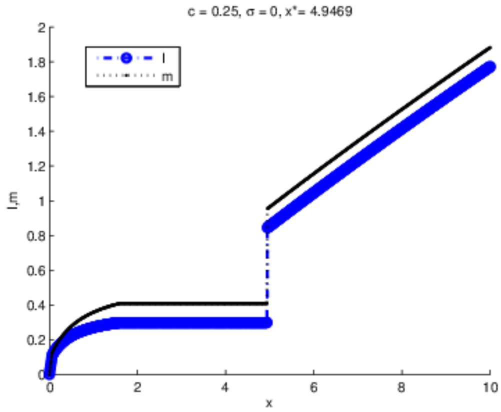


Figure 1.

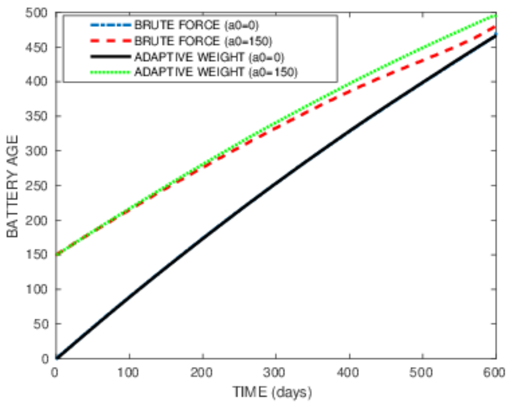


Figure 2.

7.1.3. Aircraft model identification and trajectory optimization

During the PhD of C. Rommel co-supervised with startup Safety Line, we investigated several formulations and methods for identifying an aircraft dynamics from recorded flight data. In particular, in [14] we introduce a block-sparse Bolasso approach for variable selection. In [12] we study how to quantify the closedness of a trajectory to a set of reference ones, based on the meean marginal likelihood. These works are combined with a gaussian mixture model in [15], allowing for a trade-off between optimality and acceptability of the aircraft trajectories.

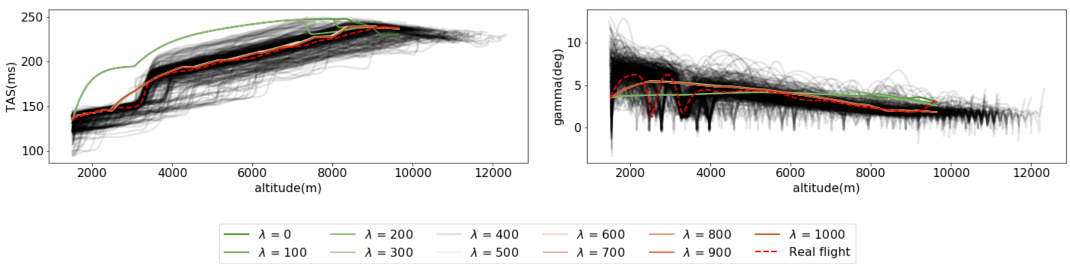


Figure 3.

Aircraft trajectory optimization. Illustration of the trade-off between performance (consumption) and acceptability (weighted by λ).

7.1.4. Microalgae cultivation in a turbid medium

In the context of IPL Algae in Silico, we study in [11] the cultivation of microalgae in a turbid medium. Microalgae cultivation with wastewater is a promising way of reducing the energetic needs for wastewater treatment and the costs of biofuel production. However, the very turbid medium is not favorable for the development of microalgae. Indeed, light, the key element for photosynthesis, rapidly vanishes along depth due to absorption and scattering. Therefore it is crucial to understand the effects of the depth on turbid cultures. In this work, we study theoretically the long-term behavior of a continuous culture of microalgae exposed to a periodic source of light. By allowing periodic variations of the depth and the hydraulic retention time, we show that the microalgae population is forced to a periodic regime. Finally, we address numerically the problem of determining the optimal variations of the depth and the hydraulic retention time for maximizing the productivity of the culture in the periodic regime.

7.1.5. Optimizing running a race on a curved track

Following on a previous study of optimal running strategies [16], we investigate in [9] the case of a curved track. In order to determine the optimal strategy to run a race on a curved track according to the lane number, we introduce a model based on differential equations for the velocity, the propulsive force and the anaerobic energy which takes into account the centrifugal force. This allows us to analyze numerically the different strategies according to the different types of track since the straight line is not always of the same length. In particular, we find that the tracks with shorter straight lines lead to better performances, while the double bend track with the longest straight line leads to the worst performances and the biggest difference between lanes. Then for a race with two runners, we introduce a psychological attraction to follow someone just ahead and the delay to benefit again from this interaction after being overtaken. We provide numerical simulations in different cases. Results are overall consistent with the IAAF rules for lanes drawing, indicating that middle lanes are the best, followed by the exterior lanes, interior lanes being the worst.

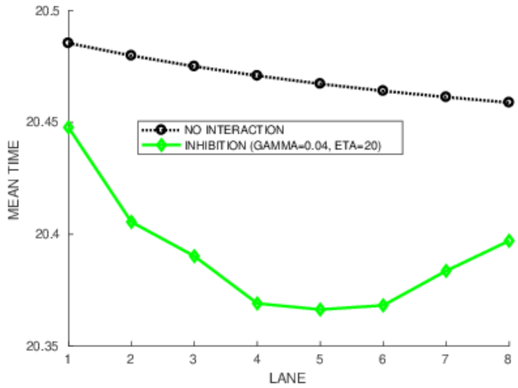


Figure 4.

Running on a curved track. Mean race times per lane, taking into account centrifugal force and psychological interaction.

7.2. Optimal control of PDEs and stochastic control

7.2.1. Sufficient optimality conditions for bilinear optimal control of the linear damped wave equation

In [8] we discuss sufficient optimality conditions for an optimal control problem for the linear damped wave equation with the damping parameter as the control. We address the case that the control enters quadratic in the cost function as well as the singular case that the control enters affine. For the non-singular case we consider strong and weak local minima , in the singular case we derive sufficient optimality conditions for weak local minima. Thereby, we take advantage of the Goh transformation applying techniques recently established in Aronna, Bonnans, and Kröner [Math. Program. 168(1):717–757, 2018]. Moreover, a numerical example for the singular case is presented.

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

In [3] we perform a variational analysis for a class of European or American options with stochastic volatility models, including those of Heston and Achdou-Tchou. Taking into account partial correlations and the presence of multiple factors, we obtain the well-posedness of the related partial differential equations, in some weigthed Sobolev spaces. This involves a generalization of the commutator analysis introduced by Achdou and Tchou.

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

In [6] we analyze an infinite horizon stochastic optimal control problem with running maximum cost. The value function is characterized as the viscosity solution of a second-order Hamilton-Jacobi-Bellman (HJB) equation with mixed boundary condition. A general numerical scheme is proposed and convergence is established under the assumptions of consistency, monotonicity and stability of the scheme. These properties are verified for a specific semi-Lagrangian scheme.

7.2.4. A stochastic data-based traffic model applied to vehicles energy consumption estimation

In the framework of the PhD of A. Le Rhun, we present in [10] a new approach to estimate traffic energy consumption via traffic data aggregation in (speed,acceleration) probability distributions. The aggregation is done on each segment composing the road network. In order to reduce data occupancy, clustering techniques are used to obtain meaningful classes of traffic conditions. Different times of the day with similar speed patterns and traffic behavior are thus grouped together in a single cluster. Different energy consumption models based on the aggregated data are proposed to estimate the energy consumption of the vehicles in the road network. For validation purposes, a microscopic traffic simulator is used to generate the data and compare the estimated energy consumption to the reference one. A thorough sensitivity analysis with respect to the parameters of the proposed method (i.e. number of clusters, size of the distributions support, etc.) is also conducted in simulation. Finally, a real-life scenario using floating car data is analyzed to evaluate the applicability and the robustness of the proposed method.

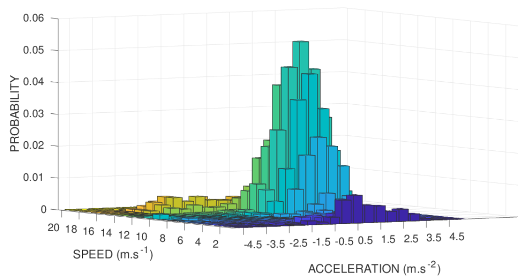


Figure 5.

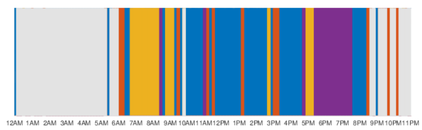


Figure 6.

Traffic modeling. Example of (speed,acceleration) distribution and illustration of clustering results with respect to day time. Slow traffic for yellow and purple clusters clearly corresponds to peak hours.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Contract with Safety Line: support of an Ilab and of a Cifre PhD. Toolbox Bocop is a component of the commercial service OptiClimb used by several airplane companies.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. IPL

9.1.1.1. Algae in Silico

Inria Project Lab ALGAE IN SILICO (2014-2018) was dedicated to provide an integrated platform for numerical simulation of microalgae “from genes to industrial process“. Commands joined the project in 2017 to tackle the optimization aspects. Our previous collaborations with teams Modemic and Biocore on bioreactors [31], [19] have been renewed in this framework, see [11]

9.1.1.2. Cosy

Inria Project Lab COSY (started in 2017) aims at exploiting the potential of state-of-art biological modelling, control techniques, synthetic biology and experimental equipment to achieve a paradigm shift in control of microbial communities. More precisely, we plan to determine and implement control strategies to make heterogeneous communities diversify and interact in the most profitable manner. Study of yeast cells has started in collaboration with team Lifeware (G. Batt) in the framework of the PhD of V. Andreani, and is pursued in the PhD of E. Weill (started 2018).

9.2. International Research Visitors

9.2.1. Visits of International Scientists

Several visits by L. Pfeiffer, U. Graz, and A. Kröner, U. Humboldt.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Selection

10.1.1.1. Member of the Conference Program Committees

- F. Bonnans: PGMO Days 2018.

10.1.2. Journal

10.1.2.1. Member of the Editorial Boards

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

10.1.3. Invited Talks

- F. Bonnans: CAESAR workshop, Palaiseau, Sept. 5-7, 2018.

10.1.4. Leadership within the Scientific Community

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

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master :
F. Bonnans: *Numerical analysis of partial differential equations arising in finance and stochastic control*, 18h, M2, Ecole Polytechnique and U. Paris 6, France.
F. Bonnans: *Optimal control*, 15h, M2, Optimization master (U. Paris-Saclay) and Ensta, France.
A. Kröner : Optimal control of partial differential equations, 20h, M2, Optimization master (U. Paris-Saclay), France.

10.2.2. Supervision

Finished PhD : C. Rommel, Data exploration for the optimization of aircraft trajectories. Started November 2015, finished Oct 2018 (CIFRE fellowship with Safety Line), F. Bonnans and P. Martinon.

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

PhD in progress : G. Bonnet, Efficient schemes for the Hamilton-Jacobi-Bellman equation. Started Oct. 2018. F. Bonnans and J.-M. Mirebeau, LMO, U. Orsay.

PhD in progress : P. Lavigne, Mathematical study of economic equilibria for renewable energy sources. Started Oct. 2018. F. Bonnans.

PhD in progress : E. Weill, Optimal control of partial differential equation systems: Application to heterogeneous cell populations. Started Oct. 2018. F. Bonnans and G. Batt, Inria and Institut Pasteur.

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

- F. Bonnans: codirection of a joint Allistene-Ancre commission (contribution to the national strategy for research), Numerics and Energy committee (2017-2018).
- F. Bonnans: Dimitrie Pompeiu Prize Committee (Academy of Romanian Scientists).
- P. Martinon is member of the CDT.

10.3.2. Articles and contents

- The work on races on curved tracks was covered in the Sciences page of Le Figaro.

10.3.3. Internal action

- The collaboration with startup Safety Line was presented to the “Journée Nationale des Nouveaux Arrivants”

11. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] M. S. Aronna, J. F. Bonnans, A. Kröner. *Optimal Control of Infinite Dimensional Bilinear Systems: Application to the Heat and Wave Equations*, in "Mathematical Programming B", January 2018, vol. 168, n^{o} 1-2, p. 717-757, https://hal.inria.fr/hal-01273496

[2] J. F. Bonnans, J. Gianatti, F. Silva. *On the time discretization of stochastic optimal control problems: the dynamic programming approach*, in "ESAIM: Control, Optimisation and Calculus of Variations", 2018, https://hal.inria.fr/hal-01474285

[3] J. F. BONNANS, A. KRÖNER.*Variational analysis for options with stochastic volatility and multiple factors*, in "SIAM Journal on Financial Mathematics", April 2018, vol. 9, n^{o} 2, p. 465-492, https://hal.inria.fr/hal-01516011

[4] B. HEYMANN, J. F. BONNANS, P. MARTINON, F. SILVA, F. LANAS, G. JIMENEZ.*Continuous Optimal Control Approaches to Microgrid Energy Management*, in "Energy Systems", January 2018, vol. 9, n^{o} 1, p. 59-77, https://hal.inria.fr/hal-01129393

[5] B. HEYMANN, P. MARTINON.*Optimal Battery Aging : an Adaptive Weights Dynamic Programming Algorithm*, in "Journal of Optimization Theory and Applications", August 2018 [*DOI :* 10.1007/S10957-018-1371-9], https://hal.inria.fr/hal-01349932

[6] A. KRÖNER, A. PICARELLI, H. ZIDANI.*Infinite Horizon Stochastic Optimal Control Problems with Running Maximum Cost*, in "SIAM Journal on Control and Optimization", 2018, vol. 56, n^{o} 5, p. 3296-3319 [*DOI :* 10.1137/17M115253X], https://hal.archives-ouvertes.fr/hal-01585766

[7] P. MARTINON, P. PICARD, A. RAJ.*On the Design of Optimal Health Insurance Contracts under Ex Post Moral Hazard*, in "Geneva Risk and Insurance Review", July 2018, vol. 43, n^{o} 2, p. 137-185 [*DOI :* 10.1057/S10713-018-0034-Y], https://hal-polytechnique.archives-ouvertes.fr/hal-01348551

Research Reports

[8] F. BETHKE, A. KRÖNER.*Sufficient optimality conditions for bilinear optimal control of the linear damped wave equation* , Humboldt Universität Berlin, 2018, https://hal.archives-ouvertes.fr/hal-01807699

Other Publications

[9] A. AFTALION, P. MARTINON.*Optimizing running a race on a curved track*, November 2018, https://arxiv.org/abs/1811.12321 - working paper or preprint, https://hal.inria.fr/hal-01936993

[10] A. LE RHUN, F. BONNANS, G. DE NUNZIO, T. LEROY, P. MARTINON.*A stochastic data-based traffic model applied to vehicles energy consumption estimation*, November 2018, working paper or preprint, https://hal.inria.fr/hal-01774621

[11] C. MARTÍNEZ, F. MAIRET, P. MARTINON, O. BERNARD.*Dynamics and control of a periodically forced microalgae culture*, October 2018, working paper or preprint, https://hal.inria.fr/hal-01891648

[12] C. ROMMEL, F. BONNANS, B. GREGORUTTI, P. MARTINON.*Quantifying the Closeness to a Set of Random Curves via the Mean Marginal Likelihood*, June 2018, working paper or preprint, https://hal.inria.fr/hal-01816407

[13] C. ROMMEL, F. BONNANS, B. GREGORUTTI, P. MARTINON.*Structured Feature Selection of Continuous Dynamical Systems for Aircraft Dynamics Identification*, December 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01965959

[14] C. ROMMEL, J. F. BONNANS, B. GREGORUTTI, P. MARTINON.*Block sparse linear models for learning structured dynamical systems in aeronautics*, June 2018, working paper or preprint, https://hal.inria.fr/hal-01816400

[15] C. ROMMEL, F. BONNANS, P. MARTINON, B. GREGORUTTI.*Gaussian Mixture Penalty for Trajectory Optimization Problems*, June 2018, working paper or preprint, https://hal.inria.fr/hal-01819749

References in notes

[16] A. AFTALION, J.-F. BONNANS.*Optimization of running strategies based on anaerobic energy and variations of velocity*, in "SIAM Journal on Applied Mathematics", 2014, vol. 74, n^o 5, p. 1615–1636

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Project-Team DATASHAPE

Understanding the shape of data

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

THEME
Algorithmics, Computer Algebra and Cryptology

Table of contents

Project-Team DATASHAPE

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

Keywords:

Computer Science and Digital Science:

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

Other Research Topics and Application Domains:

B1. - Life sciences
B2. - Health
B5. - Industry of the future
B9. - Society and Knowledge
B9.5. - Sciences

1. Team, Visitors, External Collaborators

Research Scientists

Jean-Daniel Boissonnat [Inria, Senior Researcher, HDR]
Frédéric Chazal [Team leader, Inria, Senior Researcher, HDR]
David Cohen-Steiner [Inria, Researcher]
Kunal Dutta [Inria, Starting Research Position]
Marc Glisse [Inria, Researcher]
Clément Maria [Inria, Researcher, from Feb 2018]
Steve Oudot [Inria, Researcher, HDR]
Mathijs Wintraecken [Inria, Starting Research Position]
Miroslav Kramár [Inria, Advanced Research Position]

Faculty Members

Guilherme Dias Da Fonseca [Univ d'Auvergne, Associate Professor, until Oct 2018]
Clément Levrard [Univ Denis Diderot, Associate Professor, from Sep 2018]

External Collaborators

Maksims Ovsjanikovs [Ecole polytechnique, Professor]
Bertrand Michel [Ecole Centrale de Nantes]

Technical Staff

Thomas Bonis [Inria, from Jul 2018]
Vincent Rouvreau [Inria]

PhD Students

Bertrand Beaufils [Sysnav, industrial grant]
Nicolas Berkouk [Inria]
Claire Brécheteau [Univ Paris-Sud, until Aug 2018]
Jeremy Sylvain Cochoy [Univ Paris-Sud, until Aug 2018]

Vincent Divol [Univ Paris-Sud, from Sep 2018]
Francois Godi [Univ de Nice - Sophia Antipolis, until Sep 2018]
Siargey Kachanovich [Inria]
Hariprasad Kannan [Inria, from Jul 2018 until Sep 2018]
Théo Lacombe [Ecole polytechnique]
Siddharth Pritam [Inria]
Owen Rouille [Inria, from Sep 2018]
Martin Royer [Inria, from Oct 2018 until Nov 2018]
Raphaël Tinarrage [Ecole Normale Supérieure Cachan]

Post-Doctoral Fellows

Mathieu Carrière [Inria, until Jul 2018]
Hariprasad Kannan [Inria, from Oct 2018]
Jisu Kim [Inria, from Nov 2018]
Martin Royer [Inria, from Dec 2018]

Visiting Scientists

Vincent Divol [Ecole Normale Supérieure Cachan, from Mar 2018 until Aug 2018]
Arijit Ghosh [Indian Statistical Institute, Kolkata, India, from Apr until May 2018, Dec 2018]
Ilaria Giulini [Univ Pierre et Marie Curie, until Sep 2018]
Wolfgang Polonik [UC Davis, from Sep 2018 until Oct 2018]
Hannah Santa Cruz Baur [CIMAT, from Jun 2018 until Aug 2018]
Hannah Schreiber [TU Graz, until Jul 2018]
Yitchzak Solomon [Brown University, from Sep 2018]
Ramsay Dyer [Berkeley Publishing, from Apr until May 2018, Dec 2018]
Mariette Yvinec [Retraitée, HDR]

Administrative Assistants

Stephanie Aubin [Inria, until Sep 2018]
Florence Barbara [Inria]
Sophie Honnorat [Inria, from Sep 2018]
Hélèna Kutniak [Inria, from Oct 2018]

2. Overall Objectives

2.1. Overall Objectives

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

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

3. Research Program

3.1. Algorithmic aspects of topological and geometric data analysis

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

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

3.2. Statistical aspects of topological and geometric data analysis

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

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

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

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

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

3.3. Topological approach for multimodal data processing

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

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

3.4. Experimental research and software development

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

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

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

4. Application Domains

4.1. Main application domains

Our work is mostly of a fundamental mathematical and algorithmic nature but finds a variety of applications in data analysis, e.g., in material science, biology, sensor networks, 3D shape analysis and processing, to name a few.

More specifically, DATASHAPE is working on the analysis of trajectories obtained from inertial sensors (PhD thesis of Bertrand Beaufils with Sysnav) and, more generally on the development of new TDA methods for Machine Learning and Artificial Intelligence for (multivariate) time-dependent data from various kinds of sensors in collaboration with Fujitsu.

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

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Books

- Jean-Daniel Boissonnat, Frédéric Chazal, Mariette Yvinec. *Geometric and Topological Inference*. Cambridge Texts in Applied Mathematics, vol. 57, Cambridge University Press, 2018.

5.1.2. Awards

- Mathieu Carrière was awarded the Prix de thèse solennel Thiessé de Rosemont / Schneider in Mathematics by the Chancellerie des Universités de Paris for his Ph.D. work under Steve Oudot's supervision (Ph.D. funded by ERC grant Gudhi), December 2018.

6. New Software and Platforms

6.1. GUDHI

Geometric Understanding in Higher Dimensions
KEYWORDS: Computational geometry - Topology
SCIENTIFIC DESCRIPTION: The current release of the GUDHI library includes: – Data structures to represent, construct and manipulate simplicial and cubical complexes. – Algorithms to compute simplicial complexes from point cloud data. – Algorithms to compute persistent homology and multi-field persistent homology. – Simplification methods via implicit representations.
FUNCTIONAL DESCRIPTION: The GUDHI open source library will provide the central data structures and algorithms that underly applications in geometry understanding in higher dimensions. It is intended to both help the development of new algorithmic solutions inside and outside the project, and to facilitate the transfer of results in applied fields.
NEWS OF THE YEAR: - Cover complex - Representation of persistence diagrams - Cech complex - weighted periodic 3d alpha-complex - sparse Rips complex - debian / docker / conda-forge packages

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

7. New Results

7.1. Algorithmic aspects of topological and geometric data analysis

7.1.1. DTM-based filtrations

Participants: Frédéric Chazal, Marc Glisse, Raphaël Tinarrage.

In collaboration with H. Anai, Y. Ike, H. Inakoshi and Y. Umeda of Fujitsu.

Despite strong stability properties, the persistent homology of filtrations classically used in Topological Data Analysis, such as, e.g. the Čech or Vietoris-Rips filtrations, are very sensitive to the presence of outliers in the data from which they are computed. In this paper [33], we introduce and study a new family of filtrations, the DTM-filtrations, built on top of point clouds in the Euclidean space which are more robust to noise and outliers. The approach adopted in this work relies on the notion of distance-to-measure functions, and extends some previous work on the approximation of such functions.

7.1.2. Persistent Homology with Dimensionality Reduction: k-Distance vs Gaussian Kernels

Participants: Shreya Arya, Jean-Daniel Boissonnat, Kunal Dutta.

We investigate the effectiveness of dimensionality reduction for computing the persistent homology for both k-distance and kernel distance [34]. For k-distance, we show that the standard Johnson-Lindenstrauss reduction preserves the k-distance, which preserves the persistent homology upto a $(1-\varepsilon)^{-1}$ factor with target dimension $O(k \log n/\varepsilon^2)$. We also prove a concentration inequality for sums of dependent chi-squared random variables, which, under some conditions, allows the persistent homology to be preserved in $O(\log n/\varepsilon^2)$ dimensions. This answers an open question of Sheehy. For Gaussian kernels, we show that the standard Johnson-Lindenstrauss reduction preserves the persistent homology up to an $4(1-\epsilon)^{-1}$ factor.

7.1.3. Computing Persistent Homology of Flag Complexes via Strong Collapses

Participants: Jean-Daniel Boissonnat, Siddharth Pritam.

In collaboration with Divyansh Pareek (Indian Institute of Technology Bombay, India)

We introduce a fast and memory efficient approach to compute the persistent homology (PH) of a sequence of simplicial complexes. The basic idea is to simplify the complexes of the input sequence by using strong collapses, as introduced by J. Barmak and E. Miniam [DCG (2012)], and to compute the PH of an induced sequence of reduced simplicial complexes that has the same PH as the initial one. Our approach has several salient features that distinguishes it from previous work. It is not limited to filtrations (i.e. sequences of nested simplicial subcomplexes) but works for other types of sequences like towers and zigzags. To strong collapse a simplicial complex, we only need to store the maximal simplices of the complex, not the full set of all its simplices, which saves a lot of space and time. Moreover, the complexes in the sequence can be strong collapsed independently and in parallel. Finally, we can compromize between precision and time by choosing the number of simplicial complexes of the sequence we strong collapse. As a result and as demonstrated by numerous experiments on publicly available data sets, our approach is extremely fast and memory efficient in practice [27].

7.1.4. Strong Collapse for Persistence

Participants: Jean-Daniel Boissonnat, Siddharth Pritam.

In this paper, we build on the initial success of and show that further decisive progress can be obtained if one restricts the family of simplicial complexes to flag complexes. Flag complexes are fully characterized by their graph (or 1-skeleton), the other faces being obtained by computing the cliques of the graph. Hence, a flag complex can be represented by its graph, which is a very compact representation. Flag complexes are very popular and, in particular, Vietoris-Rips complexes are by far the most widely simplicial complexes used in Topological Data Analysis. It has been shown in that the persistent homology of Vietoris-Rips filtrations can be computed very efficiently using strong collapses. However, most of the time was devoted to computing the maximal cliques of the complex prior to their strong collapse. In this paper [37], we observe that the reduced complex obtained by strong collapsing a flag complex is itself a flag complex. Moreover, this reduced complex can be computed using only the 1-skeleton (or graph) of the complex, not the set of its maximal cliques. Finally, we show how to compute the equivalent filtration of the sequence of reduced flag simplicial complexes using again only 1-skeletons. x On the theory side, we show that strong collapses of flag complexes can be computed in time $O(v^2k^2)$ where v is the number of vertices of the complex and k the maximal degree of its graph. The algorithm described in this paper has been implemented and the code will be soon released in the Gudhi library. Numerous experiments show that our method outperforms previous methods, e.g. Ripser.

7.1.5. Triangulating submanifolds: An elementary and quantified version of Whitney's method

Participants: Jean-Daniel Boissonnat, Siargey Kachanovich, Mathijs Wintraecken.

We quantize Whitney's construction to prove the existence of a triangulation for any C^2 manifold, so that we get an algorithm with explicit bounds. We also give a new elementary proof, which is completely geometric [36].

7.1.6. Randomized incremental construction of Delaunay triangulations of nice point sets

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

In collaboration with Olivier Devillers (Inria, CNRS, Loria, Université de Lorraine).

Randomized incremental construction (RIC) is one of the most important paradigms for building geometric data structures. Clarkson and Shor developed a general theory that led to numerous algorithms that are both simple and efficient in theory and in practice.

Randomized incremental constructions are most of the time space and time optimal in the worst-case, as exemplified by the construction of convex hulls, Delaunay triangulations and arrangements of line segments.

However, the worst-case scenario occurs rarely in practice and we would like to understand how RIC behaves when the input is nice in the sense that the associated output is significantly smaller than in the worst-case. For example, it is known that the Delaunay triangulations of nicely distributed points in $\mathbb{R}^d$ or on polyhedral surfaces in $\mathbb{R}^3$ has linear complexity, as opposed to a worst-case complexity of $\Theta(n^{\lfloor d/2 \rfloor})$ in the first case and quadratic in the second. The standard analysis does not provide accurate bounds on the complexity of such cases and we aim at establishing such bounds in this paper [35]. More precisely, we will show that, in the two cases above and variants of them, the complexity of the usual RIC is $O(n \log n)$, which is optimal. In other words, without any modification, RIC nicely adapts to good cases of practical value.

Along the way, we prove a probabilistic lemma for sampling without replacement, which may be of independent interest.

7.1.7. Approximate Polytope Membership Queries

Participant: Guilherme Da Fonseca.

In collaboration with Sunil Arya (Hong Kong University of Science and Technology) and David Mount (University of Maryland).

In the polytope membership problem, a convex polytope K in $\mathbb{R}^d$ is given, and the objective is to preprocess K into a data structure so that, given any query point $q \in \mathbb{R}^d$, it is possible to determine efficiently whether $q \in K$. We consider this problem in an approximate setting. Given an approximation parameter ϵ, the query can be answered either way if the distance from q to K's boundary is at most ϵ times K's diameter. We assume that the dimension d is fixed, and K is presented as the intersection of n halfspaces. Previous solutions to approximate polytope membership were based on straightforward applications of classic polytope approximation techniques by Dudley (1974) and Bentley et al. (1982). The former is optimal in the worst-case with respect to space, and the latter is optimal with respect to query time. We present four main results. First, we show how to combine the two above techniques to obtain a simple space-time trade-off. Second, we present an algorithm that dramatically improves this trade-off. In particular, for any constant $\alpha \geq 4$, this data structure achieves query time roughly $O(1/\epsilon^{(d-1)/\alpha})$ and space roughly $O(1/\epsilon^{(d-1)(1-\Omega(\log \alpha))/\alpha})$. We do not know whether this space bound is tight, but our third result shows that there is a convex body such that our algorithm achieves a space of at least $\Omega(1/\epsilon^{(d-1)(1-O(\sqrt{\alpha}))/\alpha})$. Our fourth result shows that it is possible to reduce approximate Euclidean nearest neighbor searching to approximate polytope membership queries. Combined with the above results, this provides significant improvements to the best known space-time trade-offs for approximate nearest neighbor searching in $\mathbb{R}^d$. For example, we show that it is possible to achieve a query time of roughly $O(\log n + 1/\epsilon^{d/4})$ with space roughly $O(n/\epsilon^{d/4})$, thus reducing by half the exponent in the space bound [11].

7.1.8. Approximate Convex Intersection Detection with Applications to Width and Minkowski Sums

Participant: Guilherme Da Fonseca.

In collaboration with Sunil Arya (Hong Kong University of Science and Technology) and David Mount (University of Maryland).

Approximation problems involving a single convex body in d-dimensional space have received a great deal of attention in the computational geometry community. In contrast, works involving multiple convex bodies are generally limited to dimensions $d \leq 3$ and/or do not consider approximation. In this paper, we consider approximations to two natural problems involving multiple convex bodies: detecting whether two polytopes intersect and computing their Minkowski sum. Given an approximation parameter $\epsilon > 0$, we show how to independently preprocess two polytopes A, B into data structures of size $O(1/\epsilon^{(d-1)/2})$ such that we can answer in polylogarithmic time whether A and B intersect approximately. More generally, we can answer this for the images of A and B under affine transformations. Next, we show how to ϵ-approximate the Minkowski sum of two given polytopes defined as the intersection of n halfspaces in $O(n \log(1/\epsilon) + 1/\epsilon^{(d-1)/2+\alpha})$ time, for any constant $\alpha > 0$. Finally, we present a surprising impact of these results to a well studied problem that considers a single convex body. We show how to ϵ-approximate the width of a set of n points in $O(n \log(1/\epsilon) + 1/\epsilon^{(d-1)/2+\alpha})$ time, for any constant $\alpha > 0$, a major improvement over the previous bound of roughly $O(n + 1/\epsilon^{d-1})$ time [22].

7.1.9. Approximating the Spectrum of a Graph

Participant: David Cohen-Steiner.

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

The spectrum of a network or graph $G = (V, E)$ with adjacency matrix A , consists of the eigenvalues of the normalized Laplacian $L = I - D^{-1/2}AD^{-1/2}$. This set of eigenvalues encapsulates many aspects of the structure of the graph, including the extent to which the graph posses community structures at multiple scales. We study the problem of approximating the spectrum, $\lambda = (\lambda_1, \cdots, \lambda_{|V|})$, of G in the regime where the graph is too large to explicitly calculate the spectrum. We present a sublinear time algorithm that, given the ability to query a random node in the graph and select a random neighbor of a given node, computes a succinct representation of an approximation $\widetilde{\lambda} = (\widetilde{\lambda}_1, \cdots, \widetilde{\lambda}_{|V|})$, such that $\|\widetilde{\lambda} - \lambda\|_1 \leq \varepsilon|V|$. Our algorithm has query complexity and running time $\exp(O(1/\varepsilon))$, which is independent of the size of the graph, $|V|$. We demonstrate the practical viability of our algorithm on synthetically generated graphs, and on 15 different real-world graphs from the Stanford Large Network Dataset Collection, including social networks, academic collaboration graphs, and road networks. For the smallest of these graphs, we are able to validate the accuracy of our algorithm by explicitly calculating the true spectrum; for the larger graphs, such a calculation is computationally prohibitive. The spectra of these real-world networks reveal insights into the structural similarities and differences between them, illustrating the potential value of our algorithm for efficiently approximating the spectrum of large networks [29].

7.1.10. Spectral Properties of Radial Kernels and Clustering in High Dimensions

Participants: David Cohen-Steiner, Alba Chiara de Vitis.

In this paper [40], we study the spectrum and the eigenvectors of radial kernels for mixtures of distributions in $\mathbb{R}^n$. Our approach focuses on high dimensions and relies solely on the concentration properties of the components in the mixture. We give several results describing of the structure of kernel matrices for a sample drawn from such a mixture. Based on these results, we analyze the ability of kernel PCA to cluster high dimensional mixtures. In particular, we exhibit a specific kernel leading to a simple spectral algorithm for clustering mixtures with possibly common means but different covariance matrices. This algorithm will succeed if the angle between any two covariance matrices in the mixture (seen as vectors in $\mathbb{R}^{n^2}$) is larger than $\Omega(n^{-1/6} \log^{5/3} n)$. In particular, the required angular separation tends to 0 as the dimension tends to infinity. To the best of our knowledge, this is the first polynomial time algorithm for clustering such mixtures beyond the Gaussian case.

7.1.11. Exact computation of the matching distance on 2-parameter persistence modules

Participant: Steve Oudot.

In collaboration with Michael Kerber (T.U. Graz) and Michael Lesnick (SUNY).

The matching distance is a pseudometric on multi-parameter persistence modules, defined in terms of the weighted bottleneck distance on the restriction of the modules to affine lines. It is known that this distance is stable in a reasonable sense, and can be efficiently approximated, which makes it a promising tool for practical applications. In [44] we show that in the 2-parameter setting, the matching distance can be computed exactly in polynomial time. Our approach subdivides the space of affine lines into regions, via a line arrangement. In each region, the matching distance restricts to a simple analytic function, whose maximum is easily computed. As a byproduct, our analysis establishes that the matching distance is a rational number, if the bigrades of the input modules are rational.

7.1.12. A Comparison Framework for Interleaved Persistence Modules

Participant: Miroslav Kramár.

In collaboration with Rachel Levanger (UPenn), Shaun Harker and Konstantin Mischaikow (Rutgers).

In [43], we present a generalization of the induced matching theorem of [1] and use it to prove a generalization of the algebraic stability theorem for R-indexed pointwise finite-dimensional persistence modules. Via numerous examples, we show how the generalized algebraic stability theorem enables the computation of rigorous error bounds in the space of persistence diagrams that go beyond the typical formulation in terms of bottleneck (or log bottleneck) distance.

7.1.13. Discrete Morse Theory for Computing Zigzag Persistence

Participant: Clément Maria.

In collaboration with Hannah Schreiber (Graz University of Technology, Austria)

We introduce a framework to simplify zigzag filtrations of general complexes using discrete Morse theory, in order to accelerate the computation of zigzag persistence. Zigzag persistence is a powerful algebraic generalization of persistent homology. However, its computation is much slower in practice, and the usual optimization techniques cannot be used to compute it. Our approach is different in that it preprocesses the filtration before computation. Using discrete Morse theory, we get a much smaller zigzag filtration with same persistence. The new filtration contains general complexes. We introduce new update procedures to modify on the fly the algebraic data (the zigzag persistence matrix) under the new combinatorial changes induced by the Morse reduction. Our approach is significantly faster in practice [45].

7.2. Statistical aspects of topological and geometric data analysis

7.2.1. Robust Bregman Clustering

Participants: Claire Brécheteau, Clément Levrard.

In collaboration with Aurélie Fischer (Université Paris-Diderot).

Using a trimming approach, in [38], we investigate a k-means type method based on Bregman divergences for clustering data possibly corrupted with clutter noise. The main interest of Bregman divergences is that the standard Lloyd algorithm adapts to these distortion measures, and they are well-suited for clustering data sampled according to mixture models from exponential families. We prove that there exists an optimal codebook, and that an empirically optimal codebook converges a.s. to an optimal codebook in the distortion sense. Moreover, we obtain the sub-Gaussian rate of convergence for k-means $1 \sqrt{} n$ under mild tail assumptions. Also, we derive a Lloyd-type algorithm with a trimming parameter that can be selected from data according to some heuristic, and present some experimental results.

7.2.2. Statistical analysis and parameter selection for Mapper

Participants: Mathieu Carrière, Bertrand Michel, Steve Oudot.

In [15] we study the question of the statistical convergence of the 1-dimensional Mapper to its continuous analogue, the Reeb graph. We show that the Mapper is an optimal estimator of the Reeb graph, which gives, as a byproduct, a method to automatically tune its parameters and compute confidence regions on its topological features, such as its loops and flares. This allows to circumvent the issue of testing a large grid of parameters and keeping the most stable ones in the brute-force setting, which is widely used in visualization, clustering and feature selection with the Mapper.

7.2.3. A Fuzzy Clustering Algorithm for the Mode-Seeking Framework

Participants: Thomas Bonis, Steve Oudot.

In [13] we propose a new soft clustering algorithm based on the mode-seeking framework. Given a point cloud in $\mathbb{R}^d$, we define regions of high density that we call cluster cores, then we implement a random walk on a neighborhood graph built on top of the data points. This random walk is designed in such a way that it is attracted by high-density regions, the intensity of the attraction being controlled by a temperature parameter $\beta > 0$. The membership of a point to a given cluster is then the probability for the random walk starting at this point to hit the corresponding cluster core before any other. While many properties of random walks (such as hitting times, commute distances, etc) are known to eventually encode purely local information when the number of data points grows to infinity, the regularization introduced by the use of cluster cores allows the output of our algorithm to converge to quantities involving the global structure of the underlying density function. Empirically, we show how the choice of β influences the behavior of our algorithm: for small values of β the result is really close to hard mode-seeking, while for values of β close to 1 the result is similar to the output of the (soft) spectral clustering. We also demonstrate the scalability of our approach experimentally.

7.2.4. Large Scale computation of Means and Clusters for Persistence Diagrams using Optimal Transport

Participants: Théo Lacombe, Steve Oudot.

In collaboration with Marco Cuturi (ENSAE).

Persistence diagrams (PDs) are at the core of topological data analysis. They provide succinct descriptors encoding the underlying topology of sophisticated data. PDs are backed-up by strong theoretical results regarding their stability and have been used in various learning contexts. However, they do not live in a space naturally endowed with a Hilbert structure where natural metrics are not even differentiable, thus not suited to optimization process. Therefore, basic statistical notions such as the barycenter of a finite sample of PDs are not properly defined. In [30] we provide a theoretically good and computationally tractable framework to estimate the barycenter of a set of persistence diagrams. This construction is based on the theory of Optimal Transport (OT) and endows the space of PDs with a metric inspired from regularized Wasserstein distances.

7.2.5. The k-PDTM : a coreset for robust geometric inference

Participants: Claire Brécheteau, Clément Levrard.

Analyzing the sub-level sets of the distance to a compact sub-manifold of $\mathbb{R}^d$ is a common method in TDA to understand its topology. The distance to measure (DTM) was introduced by Chazal, Cohen-Steiner and Mérigot to face the non-robustness of the distance to a compact set to noise and outliers. This function makes possible the inference of the topology of a compact subset of $\mathbb{R}^d$ from a noisy cloud of n points lying nearby in the Wasserstein sense. In practice, these sub-level sets may be computed using approximations of the DTM such as the q-witnessed distance or other power distance. These approaches lead eventually to compute the homology of unions of n growing balls, that might become intractable whenever n is large. To simultaneously face the two problems of large number of points and noise, we introduce in [39] the k-power distance to measure (k-PDTM). This new approximation of the distance to measure may be thought of as a k-coreset based approximation of the DTM. Its sublevel sets consist in union of k-balls, $k << n$, and this distance is also proved robust to noise. We assess the quality of this approximation for k possibly dramatically smaller than n, for instance $k = n13$ is proved to be optimal for 2-dimensional shapes. We also provide an algorithm to compute this k-PDTM.

7.2.6. The density of expected persistence diagrams and its kernel based estimation

Participants: Frédéric Chazal, Vincent Divol.

Persistence diagrams play a fundamental role in Topological Data Analysis where they are used as topological descriptors of filtrations built on top of data. They consist in discrete multisets of points in the plane $\mathbb{R}^2$ that can equivalently be seen as discrete measures in $\mathbb{R}^2$. When the data come as a random point cloud, these discrete measures become random measures whose expectation is studied in this paper. In [28] we first show that for a wide class of filtrations, including the Čech and Rips-Vietoris filtrations, the expected persistence diagram, that is a deterministic measure on $\mathbb{R}^2$, has a density with respect to the Lebesgue measure. Second, building on the previous result we show that the persistence surface recently introduced by Adams et al can be seen as a kernel estimator of this density. We propose a cross-validation scheme for selecting an optimal bandwidth, which is proven to be a consistent procedure to estimate the density.

7.2.7. On the choice of weight functions for linear representations of persistence diagrams

Participant: Vincent Divol.

In collaboration with Wolfgang Polonik (UC Davis)

Persistence diagrams are efficient descriptors of the topology of a point cloud. As they do not naturally belong to a Hilbert space, standard statistical methods cannot be directly applied to them. Instead, feature maps (or representations) are commonly used for the analysis. A large class of feature maps, which we call linear, depends on some weight functions, the choice of which is a critical issue. An important criterion to choose a weight function is to ensure stability of the feature maps with respect to Wasserstein distances on diagrams. In [42], we improve known results on the stability of such maps, and extend it to general weight functions. We also address the choice of the weight function by considering an asymptotic setting; assume that X_n is an i.i.d. sample from a density on $[0,1]^d$. For the Cech and Rips filtrations, we characterize the weight functions for which the corresponding feature maps converge as n approaches infinity, and by doing so, we prove laws of large numbers for the total persistence of such diagrams. Both approaches lead to the same simple heuristic for tuning weight functions: if the data lies near a d-dimensional manifold, then a sensible choice of weight function is the persistence to the power α with $\alpha \geq d$.

7.2.8. Estimating the Reach of a Manifold

Participants: Frédéric Chazal, Bertrand Michel.

In collaboration with E. Aamari (CNRS Paris 7), J.Kim, A. Rinaldo and L. Wasserman (Carnegie Mellon University).

Various problems in manifold estimation make use of a quantity called the reach, denoted by τ_M, which is a measure of the regularity of the manifold. [32] is the first investigation into the problem of how to estimate the reach. First, we study the geometry of the reach through an approximation perspective. We derive new geometric results on the reach for submanifolds without boundary. An estimator $\widehat{\tau}$ of τ_M is proposed in a framework where tangent spaces are known, and bounds assessing its efficiency are derived. In the case of i.i.d. random point cloud $\mathbb{X}_n$, $\tau(\mathbb{X}_n)$ is showed to achieve uniform expected loss bounds over a $\mathcal{C}^3$-like model. Finally, we obtain upper and lower bounds on the minimax rate for estimating the reach.

7.2.9. Robust Topological Inference: Distance To a Measure and Kernel Distance

Participants: Frédéric Chazal, Bertrand Michel.

In collaboration with B. Fasy (Univ. Montana) and F. Lecci, A. Rinaldo and L. Wasserman (Carnegie Mellon University).

Let P be a distribution with support S. The salient features of S can be quantified with persistent homology, which summarizes topological features of the sublevel sets of the distance function (the distance of any point x to S). Given a sample from P we can infer the persistent homology using an empirical version of the distance function. However, the empirical distance function is highly non-robust to noise and outliers. Even one outlier is deadly. The distance-to-a-measure (DTM), introduced by Chazal et al. (2011), and the kernel distance, introduced by Phillips et al. (2014), are smooth functions that provide useful topological information but are robust to noise and outliers. Chazal et al. (2015) derived concentration bounds for DTM. Building on these results, in [16], we derive limiting distributions and confidence sets, and we propose a method for choosing tuning parameters.

7.3. Topological approach for multimodal data processing

7.3.1. Barcode Embeddings for Metric Graphs

Participants: Steve Oudot, Yitchzak Solomon.

Stable topological invariants are a cornerstone of persistence theory and applied topology, but their discriminative properties are often poorly-understood. In [46] we study a rich homology-based invariant first defined by Dey, Shi, and Wang, which we think of as embedding a metric graph in the barcode space. We prove that this invariant is locally injective on the space of metric graphs and globally injective on a GH-dense subset. Moreover, we define a new topology on MGraphs, which we call the fibered topology, for which the barcode transform is injective on a generic (open and dense) subset.

7.3.2. Inverse Problems in Topological Persistence: a Survey

Participants: Steve Oudot, Yitchzak Solomon.

In [47] we review the literature on inverse problems in topological persistence theory. The first half of the survey is concerned with the question of surjectivity, i.e. the existence of right inverses, and the second half focuses on injectivity, i.e. left inverses. Throughout, we highlight the tools and theorems that underlie these advances, and direct the reader's attention to open problems, both theoretical and applied.

7.4. Experimental research and software development

7.4.1. Activity recognition from stride detection: a machine learning approach based on geometric patterns and trajectory reconstruction.

Participants: Bertrand Beaufils, Frédéric Chazal, Bertrand Michel.

In collaboration with M. Grelet (Sysnav).

In [23] algorithm for activity recognition is proposed using inertial sensors worn on the ankle. This innovative approach based on geometric patterns uses a stride detector that can detect both normal walking strides and atypical strides such as small steps, side steps and backward walking that existing methods struggle to detect. It is also robust in critical situations, when for example the wearer is sitting and moving the ankle, while most algorithms in the literature would wrongly detect strides. A technique inspired by Zero Velocity Update is used on the stride detection to compute the trajectory of the device. It allows to compute relevant features for the activity recognition learning task. Compared to most algorithms in the literature, this method does not use fixed-size sliding window that could be too short to provide enough information or too long and leads to overlapping issue when the window covers two different activities.

7.4.2. Dynamics of silo deformation under granular discharge

Participant: Miroslav Kramár.

In collaboration with Claudia Colonnello.

In [17], we use Topological Data Analysis to study the post buckling behavior of laboratory scale cylindrical silos under gravity driven granular discharges. Thin walled silos buckle during the discharge if the initial height of the granular column is large enough. The deformation of the silo is reversible as long as the filling height does not exceed a critical value, Lc. Beyond this threshold the deformation becomes permanent and the silo often collapses. We study the dynamics of reversible and irreversible deformation processes, varying the initial filling height around Lc. We find that all reversible processes exhibit striking similarities and they alternate between regimes of slow and fast dynamics. The patterns that occur at the beginning of irreversible deformation processes are topologically very similar to those that arise during reversible processes. However, the dynamics of reversible and irreversible processes is significantly different. In particular, the evolution of irreversible processes is much faster. This allows us to make an early prediction of the collapse of the silo based solely on observations of the deformation patterns.

7.4.3. Characterizing Granular Networks Using Topological Metrics

Participant: Miroslav Kramár.

In collaboration with Joshua Dijksman (Duke Physics), Lenka Kovalcinova and Lou Kondic (NJIT), Jie Ren (Merck Research Lab), Robert Behringer (Duke), and Konstantin Mischaikow (Rutgers).

In [18], we carry out a direct comparison of experimental and numerical realizations of the exact same granular system as it undergoes shear jamming. We adjust the numerical methods used to optimally represent the experimental settings and outcomes up to microscopic contact force dynamics. Measures presented here range form microscopic, through mesoscopic to system-wide characteristics of the system. Topological properties of the mesoscopic force networks provide a key link between mi-cro and macro scales. We report two main findings: the number of particles in the packing that have at least two contacts is a good predictor for the mechanical state of the system, regardless of strain history and packing density. All measures explored in both experiments and numerics, including stress tensor derived measures and contact numbers depend in a universal manner on the fraction of non-rattler particles, fNR. The force network topology also tends to show this universality, yet the shape of the master curve depends much more on the details of the numerical simulations. In particular we show that adding force noise to the numerical data set can significantly alter the topological features in the data. We conclude that both fNR and topological metrics are useful measures to consider when quantifying the state of a granular system.

7.5. Miscellaneous

7.5.1. On Order Types of Random Point Sets

Participant: Marc Glisse.

In collaboration with Olivier Devillers and Xavier Goaoc (Inria team Gamble) and Philippe Duchon (LaBRI, Université de Bordeaux).

Let P be a set of n random points chosen uniformly in the unit square. In this paper [41], we examine the typical resolution of the order type of P. First, we show that with high probability, P can be rounded to the grid of step $\frac{1}{n^{3+\epsilon}}$ without changing its order type. Second, we study algorithms for determining the order type of a point set in terms of the number of coordinate bits they require to know. We give an algorithm that requires on average $4n\log_2 n + O(n)$ bits to determine the order type of P, and show that any algorithm requires at least $4n\log_2 n - O(n\log\log n)$ bits. Both results extend to more general models of random point sets.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- Collaboration with Sysnav, a French SME with world leading expertise in navigation and geopositioning in extreme environments, on TDA, geometric approaches and machine learning for the analysis of movements of pedestrians and patients equipped with inetial sensors (CIFRE PhD of Bertrand Beaufils).
- Research collaboration with Fujitsu on the development of new TDA methods and tools for Machine learning and Artificial Intelligence (started in Dec 2017).

8.2. Bilateral Grants with Industry

- DATASHAPE and Sysnav have been selected for the ANR/DGA Challenge MALIN (funding: 700 kEuros) on pedestrian motion reconstruction in severe environments (without GPS access).

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

9.1.1.1. ANR ASPAG

Participant: Marc Glisse.

- Acronym : ASPAG.

- Type : ANR blanc.

- Title : Analysis and Probabilistic Simulations of Geometric Algorithms.

- Coordinator : Olivier Devillers (équipe Inria Gamble).

- Duration : 4 years from January 2018 to December 2021.

- Others Partners: Inria Gamble, LPSM, LABRI, Université de Rouen, IECL, Université du Littoral Côte d'Opale, Telecom ParisTech, Université Paris X (Modal'X), LAMA, Université de Poitiers, Université de Bourgogne.

- Abstract:

The analysis and processing of geometric data has become routine in a variety of human activities ranging from computer-aided design in manufacturing to the tracking of animal trajectories in ecology or geographic information systems in GPS navigation devices. Geometric algorithms and probabilistic geometric models are crucial to the treatment of all this geometric data, yet the current available knowledge is in various ways much too limited: many models are far from matching real data, and the analyses are not always relevant in practical contexts. One of the reasons for this state of affairs is that the breadth of expertise required is spread among different scientific communities (computational geometry, analysis of algorithms and stochastic geometry) that historically had very little interaction. The Aspag project brings together experts of these communities to address the problem of geometric data. We will more specifically work on the following three interdependent directions.

(1) Dependent point sets: One of the main issues of most models is the core assumption that the data points are independent and follow the same underlying distribution. Although this may be relevant in some contexts, the independence assumption is too strong for many applications.

(2) Simulation of geometric structures: The phenomena studied in (1) involve intricate random geometric structures subject to new models or constraints. A natural first step would be to build up our understanding and identify plausible conjectures through simulation. Perhaps surprisingly, the tools for an effective simulation of such complex geometric systems still need to be developed.

(3) Understanding geometric algorithms: the analysis of algorithm is an essential step in assessing the strengths and weaknesses of algorithmic principles, and is crucial to guide the choices made when designing a complex data processing pipeline. Any analysis must strike a balance between realism and tractability; the current analyses of many geometric algorithms are notoriously unrealistic. Aside from the purely scientific objectives, one of the main goals of Aspag is to bring the communities closer in the long term. As a consequence, the funding of the project is crucial to ensure that the members of the consortium will be able to interact on a very regular basis, a necessary condition for significant progress on the above challenges.

- See also: https://members.loria.fr/Olivier.Devillers/aspag/

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

9.2.1.1. GUDHI

Title: Algorithmic Foundations of Geometry Understanding in Higher Dimensions

Programm: FP7

Type: ERC

Duration: February 2014 - January 2019

Coordinator: Inria

Inria contact: Jean-Daniel Boissonnat.

The central goal of this proposal is to settle the algorithmic foundations of geometry understanding in dimensions higher than 3. We coin the term geometry understanding to encompass a collection of tasks including the computer representation and the approximation of geometric structures, and the inference of geometric or topological properties of sampled shapes. The need to understand geometric structures is ubiquitous in science and has become an essential part of scientific computing and data analysis. Geometry understanding is by no means limited to three dimensions. Many applications in physics, biology, and engineering require a keen understanding of the geometry of a variety of higher dimensional spaces to capture concise information from the underlying often highly nonlinear structure of data. Our approach is complementary to manifold learning techniques and aims at developing an effective theory for geometric and topological data analysis. To reach these objectives, the guiding principle will be to foster a symbiotic relationship between theory and practice, and to address fundamental research issues along three parallel advancing fronts. We will simultaneously develop mathematical approaches providing theoretical guarantees, effective algorithms that are amenable to theoretical analysis and rigorous experimental validation, and perennial software development. We will undertake the development of a high-quality open source software platform to implement the most important geometric data structures and algorithms at the heart of geometry understanding in higher dimensions. The platform will be a unique vehicle towards researchers from other fields and will serve as a basis for groundbreaking advances in scientific computing and data analysis.

9.3. International Research Visitors

9.3.1. Visits of International Scientists

- Wolfgang Polonik, UC Davis, California. Sept. and Oct. 2018. Statistical aspects of persistent homology.
- Arijit Ghosh, Indian Statistical Institute, Kolkata, India (December 2018)
- Ramsay Dyer, Berkeley Publishing (December 2018)

9.3.1.1. Internships

- Shreya Arya, BITS Pilani University, India, August-July 2018.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. Member of Organizing Committees

- F. Chazal co-organised the Tutorial "Machine Learning on Evolutionary Computation" at the IEEE World Congress on Computational Intelligence (WCCI), Rio de Janeiro, July 2018.
- J-D. Boissonnat was a member of the organization committee of the International Conference on Curves and Surfaces, Arcachon, July 2018.
- S. Oudot organized the mini-symposium on topological data analysis and learning at the International Conference on Curves and Surfaces, Arcachon, July 2018.

10.1.2. Scientific Events Selection

10.1.2.1. Member of the Conference Program Committees

- S. Oudot was a PC member of the International Symposium on Computational Geometry (SoCG), Budapest, Hungary, June 2018.
- David Cohen-Steiner was a PC member of the Symposium on Geometry Processing (SGP), Paris, France, July 2018, and of Shape Modeling International (SMI), Lisbon, Portugal, June 2018.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

Jean-Daniel Boissonnat is a member of the Editorial Board of *Journal of the ACM*, *Discrete and Computational Geometry*, *International Journal on Computational Geometry and Applications*.

Frédéric Chazal is a member of the Editorial Board of *SIAM Journal on Imaging Sciences*, *Discrete and Computational Geometry (Springer)*, *Graphical Models (Elsevier), and Journal of Applied and Computational Topology (Springer)*.

Steve Oudot is a member of the Editorial Board of *Journal of Computational Geometry*.

10.1.4. Invited Talks

Frédéric Chazal, Abel Symposium, Geiranger, Norway, June 2018.

Frédéric Chazal, Colloquium de Mathématiques, Math Dept. Amiens, October 2018.

Frédéric Chazal, AI Research Center at National Cheng-Kung University, Taiwan, May 2018.

Frédéric Chazal, National Center for High-performance Computing, Taiwan, May 2018.

Jean-Daniel Boissonnat, Hamilton Mathematics Institute, Trinity College, Dublin, Ireland, June 2018.

Steve Oudot, Workshop "Topological Data Analysis meets Symplectic Topology", Tel Aviv, Israel, May 2018.

Steve Oudot, Abel Symposium, Geiranger, Norway, June 2018.

Steve Oudot, Banff workshop on multiparameter persistence, Oaxaca, Mexico, August 2018.

Steve Oudot, ICERM, Brown University, Providence, USA, August 2018.

Steve Oudot, workshop on structural inference in high-dimensional models, Moscow, Russia, September 2018.

Clément Maria, Einstein workshop on Geometric and Topological Combinatorics, Freie Universität, Berlin, Germany, October 2018.

10.1.5. Leadership within the Scientific Community

Frédéric Chazal is co-responsible, with S. Arlot (Paris-Sud Univ.), of the "programme Maths-STIC" of the Labex Fondation Mathématique Jacques Hadamard (FMJH).

Frédéric Chazal has been a member of the Scientific council of the french "Agence pour les Mathematiques en Interaction avec l'Entreprise et la Societe (AMIES)" until Dec. 2018.

Frédéric Chazal is a member of the "Comité de pilotage" of the SIGMA group at SMAI.

Steve Oudot is co-organizing the monthly seminar on combinatorial and computational geometry at Institut Henri Poincaré.

Steve Oudot is co-head (with Luca Castelli-Aleardi) of the GT Géométrie Algorithmique within the GdR Informatique Mathématique.

Steve Oudot is a member of the program committee of the DataIA convergence institute.

10.1.6. Scientific Expertise

- Consulting collaboration for IFPEN to explore potential applications of TDA (from February 2018 to Dec. 2018).

10.1.7. Research Administration

Frédéric Chazal is a member of the Équipe de Direction at Inria Saclay.

Marc Glisse, responsable Raweb pour DataShape

Steve Oudot is vice-president of the Commission Scientifique at Inria Saclay.

Clément Maria is a member of the CDT at Inria Sophia Antipolis-Méditerranée.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master: Frédéric Chazal and Quentin Mérigot, Analyse Topologique des Données, 30h eq-TD, Université Paris-Sud, France.

Master: Jean-Daniel Boissonnat and Marc Glisse, Computational Geometry Learning, 36h eq-TD, M2, MPRI, France.

Master: Frédéric Cazals and Frédéric Chazal, Geometric Methods for Data Analysis, 30h eq-TD, M1, École Centrale Paris, France.

Master: Frédéric Chazal and Julien Tierny, Topological Data Analysis, 38h eq-TD, M2, Mathématiques, Vision, Apprentissage (MVA), ENS Paris-Saclay, France.

Master: Steve Oudot, Topological data analysis, 45h eq-TD, M1, École polytechnique, France.

Master: Steve Oudot, Data Analysis: geometry and topology in arbitrary dimensions, 24h eq-TD, M2, graduate program in Artificial Intelligence & Advanced Visual Computing, École polytechnique, France.

Undergrad-Master: Steve Oudot, preparatory course for international programming contests, 54h eq-TD, L3/M1, École polytechnique, France.

Summer School on topological data analysis and persistent homology: Steve Oudot, advanced topics, 6h eq-TD, Trento, Italy, June 2018.

Summer School on geometric data: Frédéric Chazal and Marc Glisse, Introduction to Topological Data Analysis, 9h eq-TD, Fréjus, Sept. 2018.

Winter School on Computational Geometry, Amirkabir University of Technology, Tehran, Iran. Course on Delaunay Triangulation of Manifolds, March 2018.

10.2.2. Supervision

PhD : Claire Brécheteau, Statistical aspects of distance-like functions , Defended on September 2018, Frédéric Chazal (co-advised by Pascal Massart).

PhD in progress: Bertrand Beaufils, Méthodes topologiques et apprentissage statistique pour l'actimétrie du piéton à partir de données de mouvement, started November 2016, Frédéric Chazal (co-advised by Bertrand Michel).

PhD: Jérémy Cochoy, Decomposition and stability of multidimensional persistence modules, Defended on December 10, 2018, Steve Oudot.

PhD in progress: Yitchzak Solomon, Inverse problems in topological data analysis, started September 1st, 2016, Steve Oudot (co-advised by Jeff Brock, Brown University).

PhD in progress: Nicolas Berkouk, Categorification of topological graph structures, started November 1st, 2016, Steve Oudot.

PhD in progress: Théo Lacombe, Statistics for persistence diagrams using optimal transport, started October 1st, 2017, Steve Oudot.

PhD in progress: Alba Chiara de Vitis, Concentration of measure and clustering, Jean-Daniel Boissonnat and David Cohen-Steiner.

PhD in progress: Siargey Kachanovich, Manifold reconstruction in higher dimensions, Jean-Daniel Boissonnat.

PhD in progress: Siddharth Pritam, Approximation algorithms in Computational Topology, Jean-Daniel Boissonnat.

PhD in progress: Raphaël Tinarrage, Persistence and stability of nerves in measured metric spaces for Topological Data Analysis, started September 1st, 2017, Frédéric Chazal and Marc Glisse.

PhD in progress: Vincent Divol, statistical aspects of TDA, started September 1st, 2017, Frédéric Chazal (co-advised by Pascal Massart).

PhD in progress: Owen Rouillé, September 2018, co-advised by C. Maria and J-D. Boissonnat.

10.2.3. Juries

J-D. Boissonnat was a member of the committee for the HDR defense of Aurélien Alvarez (Université d'Orléans).

F. Chazal was a member of the PhD defense committee of Jisu Kim (Carnegie Mellon University, advisors: A. Rinaldo and L. Wasserman), Claire Brécheteau (Université Paris-Saclay, advisors: F. Chazal and P. Massart), Hariprasad Kannan (Centrale-Supelec, advisor: N. Paragios), Dorian Nognen (Ecole Polytechnique, advisor: M. Ovsjanikov).

S. Oudot was a member of the Ph.D. defence committee of Tim Ophelders (T.U. Eindhoven, advisors: Bettina Speckmann and Kevin Buchin).

10.3. Popularization

10.3.1. Interventions

- Frédéric Chazal: Fujitsu Forum, "Topological Data Analysis: from academic success to industrial innovation", Tokyo, Japan, May 2018.
- Frédéric Chazal: "TDA and AI for biomedical applications", Kaohsiung MEdical Technology Expo, Taiwan, May 2018.

11. Bibliography

Major publications by the team in recent years

[1] D. ATTALI, U. BAUER, O. DEVILLERS, M. GLISSE, A. LIEUTIER.*Homological Reconstruction and Simplification in R3*, in "Computational Geometry", 2014 [*DOI :* 10.1016/J.COMGEO.2014.08.010], https://hal.archives-ouvertes.fr/hal-01132440

[2] J.-D. BOISSONNAT, R. DYER, A. GHOSH.*Delaunay Triangulation of Manifolds*, in "Foundations of Computational Mathematics", 2017, vol. 45, 38 [*DOI :* 10.1007/S10208-017-9344-1], https://hal.inria.fr/hal-01509888

[3] J.-D. BOISSONNAT, R. DYER, A. GHOSH, S. Y. OUDOT.*Only distances are required to reconstruct submanifolds*, in "Computational Geometry", 2017, vol. 66, p. 32 - 67 [*DOI :* 10.1016/J.COMGEO.2017.08.001], https://hal.inria.fr/hal-01583086

[4] J.-D. BOISSONNAT, K. C. SRIKANTA, S. TAVENAS.*Building Efficient and Compact Data Structures for Simplicial Complexe*, in "Algorithmica", September 2016 [*DOI :* 10.1007/S00453-016-0207-Y], https://hal.inria.fr/hal-01364648

[5] F. CHAZAL, D. COHEN-STEINER, A. LIEUTIER.*A Sampling Theory for Compact Sets in Euclidean Space*, in "Discrete Comput. Geom.", 2009, vol. 41, n^{o} 3, p. 461–479, http://dx.doi.org/10.1007/s00454-009-9144-8

[6] F. CHAZAL, D. COHEN-STEINER, Q. MÉRIGOT.*Geometric Inference for Measures based on Distance Functions*, in "Foundations of Computational Mathematics", 2011, vol. 11, n^{o} 6, p. 733-751, RR-6930 [*DOI :* 10.1007/S10208-011-9098-0], http://hal.inria.fr/inria-00383685

[7] F. CHAZAL, S. Y. OUDOT, M. GLISSE, V. DE SILVA.*The Structure and Stability of Persistence Modules*, SpringerBriefs in Mathematics, Springer Verlag, 2016, VII, 116, https://hal.inria.fr/hal-01330678

[8] L. J. GUIBAS, S. Y. OUDOT, P. SKRABA, F. CHAZAL.*Persistence-Based Clustering in Riemannian Manifolds*, in "Journal of the ACM", November 2013, vol. 60, n^{o} 6, 38, http://hal.archives-ouvertes.fr/hal-00923563

[9] M. MANDAD, D. COHEN-STEINER, L. KOBBELT, P. ALLIEZ, M. DESBRUN.*Variance-Minimizing Transport Plans for Inter-surface Mapping*, in "ACM Transactions on Graphics", 2017, vol. 36, 14 [*DOI :* 10.1145/3072959.3073671], https://hal.inria.fr/hal-01519006

[10] S. Y. OUDOT.*Persistence Theory: From Quiver Representations to Data Analysis*, Mathematical Surveys and Monographs, American Mathematical Society, 2015, n^{o} 209, 218, https://hal.inria.fr/hal-01247501

Publications of the year

Articles in International Peer-Reviewed Journal

[11] S. ARYA, G. DA FONSECA, D. MOUNT.*Approximate Polytope Membership Queries*, in "SIAM Journal on Computing", January 2018, vol. 47, n^{o} 1, p. 1 - 51 [*DOI :* 10.1137/16M1061096], https://hal.archives-ouvertes.fr/hal-01890054

[12] J.-D. BOISSONNAT, C. KARTHIK.*An Efficient Representation for Filtrations of Simplicial Complexes*, in "ACM Transactions on Algorithms", September 2018, vol. 14, https://hal.inria.fr/hal-01883836

[13] T. BONIS, S. OUDOT.*A Fuzzy Clustering Algorithm for the Mode-Seeking Framework*, in "Pattern Recognition Letters", 2018, https://arxiv.org/abs/1406.7130 - Submitted to Pattern Recognition Letters [*DOI :* 10.1016/J.PATREC.2017.11.019], https://hal.inria.fr/hal-01111854

[14] D. BREMNER, O. DEVILLERS, M. GLISSE, S. LAZARD, G. LIOTTA, T. MCHEDLIDZE, G. MOROZ, S. WHITESIDES, S. WISMATH.*Monotone Simultaneous Paths Embeddings in $\mathbb{R}^d$*, in "Discrete Mathematics and Theoretical Computer Science", January 2018, vol. 20, n^{o} 1, p. 1-11 [*DOI :* 10.23638/DMTCS-20-1-1], https://hal.inria.fr/hal-01529154

[15] M. CARRIERE, B. MICHEL, S. Y. OUDOT.*Statistical analysis and parameter selection for Mapper*, in "Journal of Machine Learning Research", 2018, https://hal.archives-ouvertes.fr/hal-01633106

[16] F. CHAZAL, B. FASY, F. LECCI, B. MICHEL, A. RINALDO, L. WASSERMAN.*Robust Topological Inference: Distance To a Measure and Kernel Distance*, in "Journal of Machine Learning Research", 2018, vol. 18, n^o 159, 40, https://arxiv.org/abs/1412.7197v1 - To appear, forthcoming, https://hal.inria.fr/hal-01232217

[17] C. COLONNELLO, M. KRAMÁR.*Dynamics of silo deformation under granular discharge*, in "pre physical review E", 2018, https://hal.inria.fr/hal-01849279

[18] J. DIJKSMAN, L. KOVALCINOVA, J. REN, R. BEHRINGER, M. KRAMÁR, K. MISCHAIKOW, L. KONDIC.*Characterizing Granular Networks Using Topological Metrics*, in "Physical Review E ", 2018, vol. 97, 042903, https://hal.inria.fr/hal-01757079

[19] P. DLOTKO, T. WANNER.*Rigorous cubical approximation and persistent homology of continuous functions*, in "Computers and Mathematics with Applications", March 2018, https://hal.archives-ouvertes.fr/hal-01706695

[20] L. KANARI, P. DLOTKO, M. SCOLAMIERO, R. LEVI, J. SHILLCOCK, K. HESS, H. MARKRAM.*Quantifying topological invariants of neuronal morphologies*, in "Neuroinformatics", January 2018, https://arxiv.org/abs/1603.08432 - 10 pages, 5 figures, conference or other essential info, https://hal.archives-ouvertes.fr/hal-01706793

[21] C. SORZANO, A. JIMÉNEZ, J. MOTA, J. VILAS, D. MALUENDA, M. MARTÍNEZ, E. RAMÍREZ-APORTELA, T. MAJTNER, J. SEGURA, R. SÁNCHEZ-GARCÍA, Y. RANCEL, L. DEL CAÑO, P. CONESA, R. MELERO, S. JONIC, J. VARGAS, F. CAZALS, Z. FREYBERG, J. KRIEGER, I. BAHAR, R. MARABINI, J. CARAZO.*Survey of the analysis of continuous conformational variability of biological macromolecules by electron microscopy*, in "Acta crystallographica. Section F, Structural biology communications", January 2019, vol. 75, n^o 1, p. 19-32, https://hal.inria.fr/hal-01968170

International Conferences with Proceedings

[22] S. ARYA, G. D. DA FONSECA, D. M. MOUNT.*Approximate Convex Intersection Detection with Applications to Width and Minkowski Sums*, in "ESA 2018 - European Symposium on Algorithms", Helsinki, Finland, 26th Annual European Symposium on Algorithms proceedings, August 2018, https://arxiv.org/abs/1807.00484 [*DOI :* 10.4230/LIPIcs.ESA.2018.3], https://hal.archives-ouvertes.fr/hal-01890039

[23] B. BEAUFILS, F. CHAZAL, M. GRELET, B. MICHEL.*Activity recognition from stride detection: a machine learning approach based on geometric patterns and trajectory reconstruction*, in "IPIN 2018 - 9th International Conference on Indoor Positioning and Indoor Navigation", Nantes, France, Proceedings IPIN 2018 - 9th International Conference on Indoor Positioning and Indoor Navigation, September 2018, https://hal.archives-ouvertes.fr/hal-01864467

[24] J.-D. BOISSONNAT, K. DUTTA, A. GHOSH, S. KOLAY.*Tight Kernels for Covering and Hitting: Point Hyperplane Cover and Polynomial Point Hitting Set*, in "LATIN 2018 - 13th Latin American Theoretical INformatics Symposium", Buenos Aires, Argentina, April 2018, https://hal.inria.fr/hal-01669884

[25] J.-D. BOISSONNAT, R. DYER, A. GHOSH, M. WINTRAECKEN.*Local Criteria for Triangulation of Manifolds*, in "International Symposium on Computational Geometry", Budapest, Hungary, June 2018 [*DOI :* 10.4230/LIPIcs.SoCG.2018.9], https://hal.inria.fr/hal-01801616

[26] J.-D. BOISSONNAT, A. LIEUTIER, M. WINTRAECKEN.*The reach, metric distortion, geodesic convexity and the variation of tangent spaces* , in "SoCG 2018 - 34th International Symposium on Computational Geometry", Budapest, Hungary, June 2018, p. 1-14 [*DOI :* 10.4230/LIPICS.SOCG.2018], https://hal.inria.fr/hal-01801667

[27] J.-D. BOISSONNAT, S. PRITAM, D. PAREEK.*Strong Collapse for Persistence*, in "ESA 2018 - 26th Annual European Symposium on Algorithms", Helsinki, Finland, August 2018, p. 67:1–67:13 [*DOI :* 10.4230/LIPICS], https://hal.inria.fr/hal-01886165

[28] F. CHAZAL, V. DIVOL.*The density of expected persistence diagrams and its kernel based estimation*, in "Symposium of Computational Geometry (SoCG 2018)", Budapest, Hungary, June 2018, Extended version of a paper to appear in the proceedings of the Symposium of Computational Geometry 2018, https://hal.archives-ouvertes.fr/hal-01716181

[29] D. COHEN-STEINER, W. KONG, C. SOHLER, G. VALIANT.*Approximating the Spectrum of a Graph*, in "KDD 2018 - Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining", London, United Kingdom, August 2018, https://arxiv.org/abs/1712.01725 , https://hal.inria.fr/hal-01661199

[30] T. LACOMBE, M. CUTURI, S. OUDOT.*Large Scale computation of Means and Clusters for Persistence Diagrams using Optimal Transport*, in "NIPS", Montreal, Canada, 2018, https://hal.inria.fr/hal-01966674

Scientific Books (or Scientific Book chapters)

[31] J.-D. BOISSONNAT, F. CHAZAL, M. YVINEC.*Geometric and Topological Inference*, Cambridge University Press, 2018, Cambridge Texts in Applied Mathematics, https://hal.inria.fr/hal-01615863

Other Publications

[32] E. AAMARI, J. KIM, F. CHAZAL, B. MICHEL, A. RINALDO, L. WASSERMAN.*Estimating the Reach of a Manifold*, January 2018, https://arxiv.org/abs/1705.04565 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01521955

[33] H. ANAI, F. CHAZAL, M. GLISSE, Y. IKE, H. INAKOSHI, R. TINARRAGE, Y. UMEDA.*DTM-based filtrations*, November 2018, https://arxiv.org/abs/1811.04757 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01919562

[34] S. ARYA, J.-D. BOISSONNAT, K. DUTTA.*Persistent Homology with Dimensionality Reduction: k-Distance vs Gaussian Kernels*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01950051

[35] J.-D. BOISSONNAT, O. DEVILLERS, K. DUTTA, M. GLISSE.*Randomized incremental construction of Delaunay triangulations of nice point sets*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01950119

[36] J.-D. BOISSONNAT, S. KACHANOVICH, M. WINTRAECKEN.*Triangulating submanifolds: An elementary and quantified version of Whitney's method*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01950149

[37] J.-D. BOISSONNAT, S. PRITAM.*Computing Persistent Homology of Flag Complexes via Strong Collapses*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01950074

[38] C. BRÉCHETEAU, A. FISCHER, C. LEVRARD. *Robust Bregman Clustering*, December 2018, https://arxiv.org/abs/1812.04356 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01948051

[39] C. BRÉCHETEAU, C. LEVRARD. *The k-PDTM : a coreset for robust geometric inference*, January 2018, https://arxiv.org/abs/1801.10346 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01694542

[40] D. COHEN-STEINER, A. C. DE VITIS. *Spectral Properties of Radial Kernels and Clustering in High Dimensions*, January 2019, working paper or preprint, https://hal.inria.fr/hal-01969956

[41] O. DEVILLERS, P. DUCHON, M. GLISSE, X. GOAOC. *On Order Types of Random Point Sets*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01962093

[42] V. DIVOL, W. POLONIK. *On the choice of weight functions for linear representations of persistence diagrams*, July 2018, https://arxiv.org/abs/1807.03678 - working paper or preprint, https://hal.inria.fr/hal-01833660

[43] S. HARKER, M. KRAMÁR, R. LEVANGER, K. MISCHAIKOW. *A Comparison Framework for Interleaved Persistence Modules*, April 2018, working paper or preprint, https://hal.inria.fr/hal-01757092

[44] M. KERBER, M. LESNICK, S. OUDOT. *Exact computation of the matching distance on 2-parameter persistence modules*, 2018, working paper or preprint, https://hal.inria.fr/hal-01966666

[45] C. MARIA, H. SCHREIBER. *Discrete Morse Theory for Computing Zigzag Persistence*, January 2019, https://arxiv.org/abs/1807.05172 - working paper or preprint, https://hal.inria.fr/hal-01971682

[46] S. OUDOT, E. SOLOMON. *Barcode Embeddings for Metric Graphs*, February 2018, https://arxiv.org/abs/1712.03630 - working paper or preprint, https://hal.inria.fr/hal-01708780

[47] S. OUDOT, E. SOLOMON. *Inverse Problems in Topological Persistence: a Survey*, 2018, https://arxiv.org/abs/1810.10813 - working paper or preprint, https://hal.inria.fr/hal-01966676

Project-Team DEDUCTEAM

DEDUCTEAM

IN COLLABORATION WITH: Laboratoire specification et vérification (LSV)

IN PARTNERSHIP WITH:
ENS Paris-Saclay

RESEARCH CENTER
Saclay - Île-de-France

THEME
Proofs and Verification

Table of contents

Project-Team DEDUCTEAM

Creation of the Team: 2011 December 01, updated into Project-Team: 2017 January 01

Keywords:

Computer Science and Digital Science:

A2.1.4. - Functional programming
A2.1.11. - Proof languages
A2.4.3. - Proofs
A3.1.1. - Modeling, representation
A7. - Theory of computation
A7.2. - Logic in Computer Science

Other Research Topics and Application Domains:

B7. - Transport and logistics

1. Team, Visitors, External Collaborators

Research Scientists

Gilles Dowek [Team leader, Inria, Senior Researcher, HDR]
Bruno Barras [Inria, Researcher]
Frédéric Blanqui [Inria, Researcher, HDR]

Faculty Member

Guillaume Burel [ENSIIE, Associate professor, En délégation]

External Collaborators

Jean-Pierre Jouannaud [Emeritus, HDR]
Catherine Dubois [ENSIIE, HDR]
Olivier Hermant [École Nationale Supérieure des Mines de Paris, HDR]

PhD Students

Guillaume Bury [Université Paris Diderot, until Sep 2018]
Mohamed Yacine El Haddad [Université de Paris-Saclay]
Gaspard Férey [Univserité de Paris-Saclay]
Guillaume Genestier [Université de Paris-Saclay]
François Thiré [Université de Paris-Saclay]

Post-Doctoral Fellows

Rodolphe Lepigre [Inria]
Franck Slama [Inria, from May 2018]

Administrative Assistants

Adeline Lochet [Inria, from Jul 2018]
Emmanuelle Perrot [Inria, until Jun 2018]

2. Overall Objectives

2.1. Objectives

The project-team investigates the design of logical frameworks, in order to ensure interoperability between proof systems, and to the development of system-independent proof libraries. To achieve these goals, we develop

- a logical framework DEDUKTI, where several theories can be expressed,
- tools to import proofs developed in external proof systems to DEDUKTI theories,
- tools to translate proofs from one DEDUKTI theory to another,
- tools to export proofs expressed in DEDUKTI theories to an external proof system,
- tools to prove the confluence, the termination, and the consistency of theories expressed in DEDUKTI,
- tools to develop proofs directly in DEDUKTI,
- an encyclopedia LOGIPEDIA of proofs expressed in various DEDUKTI theories.

2.2. History

The idea that systems such as Euclidean geometry or set theory should be expressed, not as independent systems, but in a logical framework appeared with the design of the first logical framework: predicate logic, in 1928. Later, several more powerful logical frameworks have been designed: λ-prolog, Isabelle, the Edinburgh logical framework, Pure type systems, and Deduction modulo theory.

The logical framework that we use is a simple λ-calculus with dependent types and rewrite rules, called the $\lambda\Pi$-calculus modulo theory, and also the Martin-Löf logical framework, and it generalizes all the mentioned frameworks. It is implemented in the system DEDUKTI.

The first version of DEDUKTI was developed in 2011 by Mathieu Boespflug [29]. From 2012 to 2015, new versions of DEDUKTI were developed and several theories were expressed in DEDUKTI, allowing to import proofs developed in MATITA (with the tool KRAJONO), HOL LIGHT (with the tool HOLIDE), FOCALIZE (with the tool FOCALIDE), IPROVER, and ZENON, totalizing several hundred of megabytes of proofs.

From 2015 to 2018, we focused on the translation of proofs from one DEDUKTI theory to another and to the exporting of proofs to other proof systems. In particular the MATITA arithmetic library has been translated to a much weaker theory: constructive simple type theory, allowing to export it to COQ, LEAN, PVS, HOL LIGHT, and ISABELLE/HOL. This led us to develop, in 2018, an online proof encyclopedia LOGIPEDIA, allowing to share and browse this library. We also focused on the development of new theories in DEDUKTI, and on an interactive theorem prover on top of DEDUKTI.

3. Research Program

3.1. Logical Frameworks

A thesis, which is at the root of our research effort, is that logical systems should be expressed as theories in a logical framework. As a consequence, proof-checking systems should not be focused on one theory, such as Simple type theory, Martin-Löf's type theory, or the Calculus of constructions, but should be theory independent. On the more theoretical side, the proof search algorithms, or the algorithmic interpretation of proofs should not depend on the theory in which proofs are expressed, but this theory should just be a parameter. This is for instance expressed in the title of our invited talk at ICALP 2012: *A theory independent Curry-De Bruijn-Howard correspondence* [31].

Various limits of Predicate logic have led to the development of various families of logical frameworks: λ-prolog and Isabelle have allowed terms containing free variables, the Edinburgh logical framework has allowed proofs to be expressed as λ-terms, Pure type systems have allowed propositions to be considered as terms, and Deduction modulo theory has allowed theories to be defined not only with axioms, but also with computation rules.

The $\lambda\Pi$-calculus modulo theory, that is implemented in the system DEDUKTI and that is a synthesis of the Edinburgh logical framework and of Deduction modulo theory, subsumes them all. Part of our research effort is focused on improving the $\lambda\Pi$-calculus modulo theory, for instance allowing to define congruences with associative and commutative rewriting. Another part of our research effort is focused on the automatic analysis of theories to prove their confluence, termination, and consistency either by pencil and paper proofs or automatically [4].

3.2. Interoperability and proof encyclopediae

Using a single prover to check proofs coming from different systems naturally leads to investigate how these proofs can be translated from one theory to another and used in a system different from the system in which they have been developed. This issue is of prime importance because developments in proof systems are getting bigger and, unlike other communities in computer science, the proof checking community has given little effort in the direction of standardization and interoperability.

For each proof, independently of the system in which it has been developed, we should be able to identify the systems in which it can be expressed. For instance, we have shown that many proofs developed in the MATITA prover did not use the full strength of the logic of MATITA and could be exported, for instance, to the systems of the HOL family, that are based on a weaker logic.

Rather than importing proofs from one system, transforming them, and exporting them to another system, we can use the same tools to develop system-independent proof encyclopedia. In such a library, each proof is labeled with the theories in which it can be expressed and so with the systems in which it can be used.

3.3. Interactive theorem proving

If our main goal with DEDUKTI is to import, transform, and export proofs developed in other systems, we also want to investigate how DEDUKTI can be used as the basis of an interactive theorem prover. This leads to two new scientific questions: first, how much can a tactic system be theory independent, and then how does rewriting extends the possibility to write tactics.

This has led to the development of a new version of DEDUKTI, which supports metavariables. Several tactics have been developed for this system, which are intended to help a human user to write proofs in our system instead of writing proof terms by hand. This work is a continuation of the previous work the team did on DEMON, which was an extension of DEDUKTI, whereas the support for interactive theorem proving is now native in DEDUKTI.

4. Application Domains

4.1. Interoperability

Our main impact applications, for instance to proofs of programs, or to air traffic control, are through our cooperation with other teams.

As a matter of fact, we view our work on interoperability and on the design of a formal proof encyclopedia as a service to the formal proof community.

5. Highlights of the Year

5.1. Highlights of the Year

Logipedia

We have launched in September the first system independent encyclopedia of formal proofs: LOGIPEDIA.

Awards

Serge Abiteboul and Gilles Dowek have received the Award *La science se livre* in January.

6. New Software and Platforms

6.1. Autotheo

KEYWORD: Automated deduction
SCIENTIFIC DESCRIPTION: Transformation of axiomatic theories into rewriting systems that can be used by iProverModulo.
FUNCTIONAL DESCRIPTION: Autotheo is a tool that transforms axiomatic theories into polarized rewriting systems, thus making them usable in iProverModulo. It supports several strategies to orient the axioms, some of them being proved to be complete, in the sense that ordered polarized resolution modulo the resulting systems is refutationally complete, some others being merely heuristics. In practice, Autotheo takes a TPTP input file and produces an input file for iProverModulo.
NEWS OF THE YEAR: Maintenance.

- Participant: Guillaume Burel
- Partner: ENSIIE
- Contact: Guillaume Burel
- Publication: Consistency Implies Cut Admissibility
- URL: http://www.ensiie.fr/~guillaume.burel/blackandwhite_autotheo.html.en

6.2. CoLoR

Coq Library on Rewriting and termination
KEYWORDS: Coq - Formalisation
FUNCTIONAL DESCRIPTION: CoLoR is a Coq library on rewriting theory and termination. It provides many definitions and theorems on various mathematical structures (quasi-ordered sets, relations, ordered semi-rings, etc.), data structures (lists, vectors, matrices, polynomials, finite graphs), term structures (strings, first-order terms, lambda-terms, etc.), transformation techniques (dependency pairs, semantic labeling, etc.) and (non-)termination criteria (polynomial and matrix interpretations, recursive path ordering, computability closure, etc.).

- Authors: Frédéric Blanqui and Sébastien Hinderer
- Contact: Frédéric Blanqui
- Publications: CoLoR: a Coq library on well-founded rewrite relations and its application to the automated verification of termination certificates - Automated Verification of Termination Certificates - CoLoR: a Coq library on rewriting and termination
- URL: http://color.inria.fr/

6.3. Coqine

Coq In dEdukti

KEYWORDS: Higher-order logic - Formal methods - Proof
FUNCTIONAL DESCRIPTION: CoqInE is a plugin for the Coq software translating Coq proofs into Dedukti terms. It provides a Dedukti signature file faithfully encoding the underlying theory of Coq (or a sufficiently large subset of it). Current development is mostly focused on implementing support for Coq universe polymorphism. The generated ouput is meant to be type-checkable using the latest version of Dedukti.

- Contact: Guillaume Burel
- URL: http://www.ensiie.fr/~guillaume.burel/blackandwhite_coqInE.html.en

6.4. Dedukti

KEYWORD: Logical Framework
FUNCTIONAL DESCRIPTION: Dedukti is a proof-checker for the LambdaPi-calculus modulo. As it can be parametrized by an arbitrary set of rewrite rules, defining an equivalence relation, this calculus can express many different theories. Dedukti has been created for this purpose: to allow the interoperability of different theories.

Dedukti's core is based on the standard algorithm for type-checking semi-full pure type systems and implements a state-of-the-art reduction machine inspired from Matita's and modified to deal with rewrite rules.

Dedukti's input language features term declarations and definitions (opaque or not) and rewrite rule definitions. A basic module system allows the user to organize his project in different files and compile them separately.

Dedukti features matching modulo beta for a large class of patterns called Miller's patterns, allowing for more rewriting rules to be implemented in Dedukti.
NEWS OF THE YEAR: There has been a new release 2.6 in 2018. This release provides a better control on module loading, and a better log of rewrite steps.

- Participants: François Thiré, Gaspard Ferey, Guillaume Genestier and Rodolphe Lepigre
- Contact: François Thiré
- Publications: Dedukti:un vérificateur de preuves universel - Rewriting Modulo β in the $\lambda\Pi$-Calculus Modulo - Expressing theories in the $\lambda\Pi$-calculus modulo theory and in the Dedukti system
- URL: https://deducteam.github.io/

6.5. Holide

KEYWORD: Proof
FUNCTIONAL DESCRIPTION: Holide translates HOL proofs to Dedukti[OT] proofs, using the OpenTheory standard (common to HOL Light and HOL4). Dedukti[OT] being the encoding of OpenTheory in Dedukti.

- Contact: Guillaume Burel
- URL: http://deducteam.gforge.inria.fr/holide/

6.6. HOT

Higher-Order Termination
FUNCTIONAL DESCRIPTION: HOT is an automated termination prover for higher-order rewriting, based on the notion of computability closure.

- Contact: Frédéric Blanqui
- URL: http://rewriting.gforge.inria.fr/hot.html

6.7. iProver Modulo

KEYWORDS: Automated deduction - Automated theorem proving
SCIENTIFIC DESCRIPTION: Integration of ordered polarized resolution modulo theory into the prover iProver.

Functional Description: iProver Modulo is an extension of the automated theorem prover iProver originally developed by Konstantin Korovin at the University of Manchester. It implements ordered polarized resolution modulo theory, a refinement of the resolution method based on deduction modulo theory. It takes as input a proposition in predicate logic and a clausal rewriting system defining the theory in which the formula has to be proved. Normalization with respect to the term rewriting rules is performed very efficiently through translation into OCaml code, compilation and dynamic linking. Experiments have shown that ordered polarized resolution modulo dramatically improves proof search compared to using raw axioms.
News Of The Year: Maintenance of Dedukti output

- Participant: Guillaume Burel
- Partner: ENSIIE
- Contact: Guillaume Burel
- Publications: A Shallow Embedding of Resolution and Superposition Proofs into the ??-Calculus Modulo - Experimenting with deduction modulo
- URL: https://github.com/gburel/iProverModulo

6.8. mSAT

Keyword: Propositional logic
Functional Description: mSAT is a modular, proof-producing, SAT and SMT core based on Alt-Ergo Zero, written in OCaml. The solver accepts user-defined terms, formulas and theory, making it a good tool for experimenting. This tool produces resolution proofs as trees in which the leaves are user-defined proof of lemmas.

- Contact: Guillaume Bury
- Publication: mSAT:An OCaml SAT Solver
- URL: https://github.com/Gbury/mSAT

6.9. Rainbow

Termination certificate verifier
Keywords: Demonstration - Code generation - Verification
Functional Description: Rainbow is a set of tools for automatically verifying the correctness of termination certificates expressed in the CPF format used in the annual international competition of termination tools. It contains: a tool xsd2coq for generating Coq data types for representing XML files valid with respect to some XML Schema, a tool xsd2ml for generating OCaml data types and functions for parsing XML files valid with respect to some XML Schema, a tool for translating a CPF file into a Coq script, and a standalone Coq certified tool for verifying the correctness of a CPF file.

- Author: Frédéric Blanqui
- Contact: Frédéric Blanqui
- Publications: Automated verification of termination certificates - Automated verification of termination certificates
- URL: http://color.inria.fr/rainbow.html

6.10. Krajono

Keyword: Proof
Functional Description: Krajono translates Matita proofs into Dedukti[CiC] (encoding of CiC in Dedukti) terms.

- Contact: François Thiré

6.11. archsat

KEYWORDS: Automated theorem proving - First-order logic - Propositional logic
FUNCTIONAL DESCRIPTION: Archsat is an automated theorem prover aimed at studying the integration of first-order theorem prover technologies, such as rewriting, into SMT solvers.

- Contact: Guillaume Bury
- URL: https://gforge.inria.fr/projects/archsat

6.12. lrat2dk

KEYWORDS: Automated theorem proving - Proof
FUNCTIONAL DESCRIPTION: Take as input a SAT proof trace in LRAT format, which can be obtained from the de facto standard format DRAT using drat-trim. Output a proof checkable by Dedukti, in a shallow encoding of propositional logic.

- Participant: Guillaume Burel
- Partner: ENSIIE
- Contact: Guillaume Burel
- URL: https://github.com/gburel/lrat2dk

6.13. ekstrakto

KEYWORDS: TPTP - TSTP - Proof assistant - Dedukti
FUNCTIONAL DESCRIPTION: Extracting TPTP problems from a TSTP trace. Proof reconstruction in Dedukti from TSTP trace.

- Contact: Mohamed Yacine El Haddad
- URL: https://github.com/elhaddadyacine/ekstrakto

6.14. SizeChangeTool

KEYWORDS: Rewriting systems - Proof assistant - Termination
FUNCTIONAL DESCRIPTION: A termination-checker for higher-order rewriting with dependent types. Took part in the Termination Competition 2018 (http://termination-portal.org/wiki/Termination_Competition_2018) in the "Higher-Order Rewriting (union Beta)" category.

- Partner: Mines ParisTech
- Contact: Guillaume Genestier
- URL: https://github.com/Deducteam/SizeChangeTool

7. New Results

7.1. $\lambda\Pi$-calculus modulo theory

Gilles Dowek, Jean-Pierre Jouannaud and Jiaxiang Liu have started a program for developing new techniques for proving confluence of dependently typed theories, which do not rely on termination. These results have been presented at Types 2016, and will be submitted to a Journal early 2019. Target applications for these techniques are encodings of the Calculus of inductive constructions with polymorphic universes in the $\lambda\Pi$-calculus modulo theory.

Frédéric Blanqui has published in the Journal of Functional Programming a long article synthesizing his work on the use of size annotations for proving termination [12]. This paper provides a general and modular criterion for the termination of simply-typed λ-calculus extended with function symbols defined by user-defined rewrite rules. Following a work of Hughes, Pareto and Sabry, for functions defined with a fixpoint operator and pattern-matching, several criteria use typing rules for bounding the height of arguments in function calls. In this paper, we extend this approach to rewriting-based function definitions and more general user-defined notions of size.

Size-change termination is a technique introduced for first-order functional programs. In [16], Frédéric Blanqui and Guillaume Genestier show how it can be used to study the termination of higher-order rewriting in the $\lambda\Pi$-calculus modulo theory.

Dependency pairs are a key concept at the core of modern automated termination provers for first-order term rewrite systems. In [22], Frédéric Blanqui, Guillaume Genestier and Olivier Hermant introduced an extension of this technique for a large class of dependently-typed higher-order rewrite systems. This improves previous results by Wahlstedt on one hand and Frédéric Blanqui on the other hand to strong normalization and non-orthogonal rewrite systems. This new criterion has been implemented in the type-checker DEDUKTI.

7.2. Dedukti

Frédéric Blanqui and Guillaume Genestier have formally defined the operational semantics of DEDUKTI 2.5, showing some problems with non left-linear rewrite rules.

Rodolphe Lepigre, Frédéric Blanqui and Franck Slama developed a new version of DEDUKTI, available on https://github.com/Deducteam/lambdapi, with meta-variables and a small set of tactics in order to be able to build DEDUKTI proofs interactively.

Aristomenis-Dionysios Papadopoulos has added a rewrite tactic in the style of Ssreflect [27].

Emilio Gallego added an LSP server for communicating with editors.

Ismail Lachheb has developed a plugin for DEDUKTI based on the LSP protocol into the Atom editor [25].

Guillaume Burel added support for polarized Deduction modulo theory in DEDUKTI.

Quentin Ye has developed an algorithm to compare λ-terms. The main point was to take sharing into account, so as to relate the complexity with the space used to represent the term, rather than with the size of the term. He has implemented this algorithm in the DEDUKTI codebase. He has also run his algorithm on examples that show an exponential speed-up compared to the naive algorithm [21].

7.3. Theories

Gaspard Férey and François Thiré defined a new encoding for Cumulative type systems (CTS) in the $\lambda\Pi$-calculus modulo theory, extending the work of Ali Assaf's PhD [28]. This encoding relies on explicit subtyping which requires additional computational rules. It provides a way to encode a larger class of CTS, which sheds a new light on the computational content of explicit subtyping. This encoding should be extendable to express more advanced features such as universe polymorphism in the Calculus of Inductive Construction, a first step to have a faithful encoding of the COQ system. The encoding has been proven correct under the hypothesis that the computational rules are confluent.

François Thiré redesigned the tool UNIVERSO, so that it can be used for a larger class of CTS. The specification for UNIVERSO can be given by rewrite rules which makes UNIVERSO much easier to use. This tool is a first step to have an automatic chain of translations to translate proofs in the encoding of MATITA to $\mathrm{STT}\forall_{\beta\delta}$.which would make these proofs interoperable with 5 different systems.

François Thiré changed the encoding provided by KRAJONO to integrate some ideas of the encoding discussed above. This encoding is compatible with the tool UNIVERSO.

Gaspard Férey updated the COQINE software to translate COQ's 8.8 version. In this version, the standard library relies on universe polymorphism so partial support for the translation of this feature was integrated. Since encodings of the many features of Coq (inductive constructions, floating universes, several kinds of universe polymorphisms, etc) are a current work in progress, the software was made parameterizable to allow experimentations of multiple encodings of these features.

Gaspard Férey showcased an encoding of the Calculus of Inductive Constructions (CiC) relying on associative-commutative (AC) rewriting on the arithmetic library translated from MATITA. This practical experiment shows the limitations of AC-rewriting (as implemented in DEDUKTI) in terms of performance and the need for special care when defining encodings relying on this feature.

Guillaume Burel began to write a tool translating SAT proof traces in LRAT format into DEDUKTI proofs. The main issue was that steps in LRAT traces are not logical consequences of previous clauses but only preserve provability.

Mohamed Yacine El Haddad developed a tool to extract TPTP problems from a TSTP trace (generated by automated theorem provers) and reconstruct the proof of the trace in DEDUKTI format.

Bruno Barras has started to develop a model of Homotopy Type Theory (HoTT) in DEDUKTI. This is basically a presheaf model, where the choice of the base category leads either to the simplicial sets model or to the cubical model of HoTT. This construction generalizes the setoid model construction [2] to an arbitrary dimension. Since this involves encoding notions of category theory, the rewriting feature of DEDUKTI is intensively used to represent, among others, the associativity of morphism composition, or the naturality conditions.

Guillaume Bury has proposed an automation-friendly set theory for the B method. This theory is expressed using first order logic extended to polymorphic types and rewriting. Rewriting is introduced along the lines of deduction modulo theory, where axioms are turned into rewrite rules over both propositions and terms. This work has been published in [30].

7.4. Interoperability

François Thiré has defined in DEDUKTI a constructive version of simple type theory with prenex polymorphism: $\mathrm{STT}\forall_{\beta\delta}$. This work has been published at the LFMTP workshop in [15]. $\mathrm{STT}\forall_{\beta\delta}$ has been used to encode an arithmetic library able to prove little Fermat's theorem. Then these proofs has been exported to different systems that are: COQ, MATITA, LEAN and OPENTHEORY. Gilles Dowek, César Muñoz, and François Thiré have developed a translation of $\mathrm{STT}\forall_{\beta\delta}$ to PVS.

Then, Walid Moustaoui and François Thiré have built a website called LOGIPEDIA which allows the user to inspect this arithmetic library and the user can download the proof of this theorem to one of the systems mentioned above.

7.5. Drags

Shared and cyclic structures are very common in both programming and proving, which requires generalizing term rewriting techniques to graphs. Jean-Pierre Jouannaud and Nachum Dershowitz have introduced a very general class of multigraphs, called drags, equipped with a composition operator $\otimes$ which provides with a rich categorical structure. Rewriting a drag D can then be defined in a very simple way, by writing D as the composition of a left-hand side of rules L and a context C, and then replacing L by R, the right-hand side of the rule, which yields the rewritten drag $R \otimes C$. The fundamental aspects of the algebra of drags have been presented at TERMGRAPH'2018 and have also been submitted to a special issue of TCS. Termination of drag rewriting in investigated in [20].

7.6. SCTL

Gilles Dowek, Liu Jian, and Ying Jiang have reworked the presentation of CTL in sequent calculus proposed by Gilles Dowek and Ying Jiang in 2012 and provided an implementation of it. This work has been published in [13].

8. Partnerships and Cooperations

8.1. National Initiatives

The ANR PROGRAMme is an ANR for junior researcher Liesbeth Demol (CNRS, UMR 8163 STL, University Lille 3) to which G. Dowek participates. The subject is: “What is a program? Historical and Philosophical perspectives”. This project aims at developing the first coherent analysis and pluralistic understanding of “program” and its implications to theory and practice.

8.2. International Initiatives

Brazil: STIC Amsud.

Argentina: Ecos

China: Inria-NSFC

8.3. Informal International Partners

Our main international partners are Alejandro Diáz-Caro (Buenos Aires), Bruno Lopes (Niteroi), Ying Jiang (Beijing), Florian Rabe (Bremen), Brigitte Pientka (McGill), César Muñoz (NASA), and Stéphane Graham-Lengrand (SRI).

8.4. International Research Visitors

Alejandro Díaz-Caro (Buenos Aires) has visited Deducteam for two weeks.

Ying Jiang (Beijing) has visited Deducteam for three weeks.

Aristomenis-Dionysios Papadopoulos (Imperial College, London) has visited Deducteam. He worked with Frédéric Blanqui on the development of a rewrite tactic in DEDUKTI [27].

8.4.1. Visits to International Teams

Gilles Dowek has spent two weeks at the University of Buenos Aires.

Gilles Dowek has spent two weeks at the Institute of Aerospace (USA).

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organization

Guillaume Burel has been local organizer of the scientific days of the CNRS GDR GPL working groups LTP and MTV2.

9.1.2. Scientific Events Selection

Frédéric Blanqui has been PC chair of the 13th International Workshop on Logical Frameworks and Meta-Languages: Theory and Practice (LFMTP’18) with Giselle Reis.

Frédéric Blanqui is Workshop Chair of LICS and member of the Steering Committee of LICS.

Frédéric Blanqui is member of the Steering Committee of the International School on Rewriting (ISR) of the WG 1.6 of the International Federation for Information Processing.

Gilles Dowek has been a PC member of TYPES 2018.

Guillaume Burel has been PC member of the 30th Journées Francophones des Langages Applicatifs.

Guillaume Burel has reviewed a submission for the International Conference on Principles and Practice of Constraint Programming (CP). Guillaume Genestier reviewed submissions to the conferences Logic in Computer Science (LICS), Principles and Practice of Declarative Programming (PPDP) and European Symposium on Programming (ESOP).

9.1.3. Journals

Gilles Dowek is an editor of TCS-C.

Frédéric Blanqui has reviewed a paper for Mathematical Structures in Computer Science (MSCS). Guillaume Burel has reviewed papers for the Computer Journal and Logical Methods in Computer Science (LMCS). Rodolphe Lepigre has reviewed a paper for International Conference on Foundations of Software Science and Computation Structures (FoSSaCS). Rodolphe Lepigre has reviewed a paper for the journal ACM Transactions on Programming Languages and Systems (TOPLAS). Franck Slama has reviewed a paper for the Journal of Functional Programming.

9.1.4. Invited Talks

- Rodolphe Lepigre gave an invited talk entitled “The PML Language: Realizability at the Service of Program Proofs” at the Realizability Workshop (12-13 June 2018) in Luminy.
- Rodolphe Lepigre gave an invited talk entitled “An Overview of the PML_2 Language: Realizability, Subtyping and Cyclic Proofs” at LRI, for the starting days of the new Scalp working group of GDR IM. This is a presentation of his paper [32].
- Gilles Dowek has given an invited talk at NFM (Nasa Formal Methods).
- Jean-Pierre Jouannaud has given an invited talk at the workshop "Rewriting Techniques for Program Transformation and Evaluation" at FLoC, on July 8, 2018.

9.1.5. Seminars

- Gilles Dowek has participated to the meeting "From Information to Cells" organized by Hélène Kirchner and Antoine Danchin. He has given a talk at the National Institute of Aerospace.
- Gilles Dowek has co-organized a seminar on Logic and Philosophy at the CNFHPST.
- Guillaume Burel has presented a talk entitled “Bridging holes on DEDUKTI proofs, an overview” at the scientific day of the Digicosme working group UPSCaLe.
- Bruno Barras has given a talk entitled “An analysis of bindlib” at the UPSCaLe meeting (June’18) held in Palaiseau.
- Mohamed Yacine EL HADDAD has presented his work at internal laboratory seminar of LSV (June’18) and SAMOVAR (November’18).
- Gaspard Férey has presented his work at internal laboratory seminar of LSV (June’18).
- Guillaume Genestier has presented his work at the internal laboratory seminar of Centre de Recherche en Informatique of Mines ParisTech (February’18) and LSV (June’18) and presented DEDUKTI at the doctoral seminar of La Société Informatique de France (June’18). He presented [16] in the WorkShop on Termination (WST) at Oxford (July’18).
- Rodolphe Lepigre has presented his work on "Termination checking using well-founded typing derivations" at a Deducteam seminar in September 2018.
- Rodolphe Lepigre has given a talk entitled “The PML_2 Language, Integrated Program Verification in ML” at the Max Planck Institute for Software Systems in Saarbrücken, in November 2018.
- Franck Slama has presented some previous work at an internal laboratory seminar of LSV in December 2017.
- François Thiré has presented his work on interoperability at the UPSCaLe seminar on March 2018, then he presented his paper [15] at the LFMT Workshop at Oxford (July’17).
- Aristomenis Papadopoulos has presented the work he did during his summer internship at a Deducteam seminar in September 2018.

9.1.6. Leadership within the scientific community

Gilles Dowek is president of the scientific board of the Société informatique de France.

He is a member of the Ethic council CERNA.

He is a member of the Comité National Français d'Histoire et de Philosophie des Sciences et des techniques.

He is a member of the scientific board of La Main à la pâte.

He is a member of the scientific board of the Institut Villebon Charpak.

He is a member of the scientific board of the Maison des sciences de Lorraine.

He is the president of the Board of teacher school (ESPE) of the University of Lorraine.

He is a member of the scientific board of SystemX.

He is a member of the scientific board of the team Humanités numériques at the Collège des Bernardins.

Gilles Dowek and Jean-Pierre Jouannaud are honorary members of IFIP-WG1.6.

Jean-Pierre Jouannaud is a permanent member of the visiting committee of Academia Sinica, Taiwan.

9.1.7. Scientific Expertise

Frédéric Blanqui reviewed a project for the Netherlands Organization for Scientific Research (NWO).

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

- Master: Bruno Barras, proof assistants, 12h, M2, MPRI
- Master: Frédéric Blanqui, formal languages, 21h, M1, ENSIIE
- Master: Frédéric Blanqui, rewriting theory, 14h, M1, ENS Paris-Saclay
- Master: Frédéric Blanqui, λ-calculus and theories in first-order logic, 18h, M1/M2, ENS Paris-Saclay
- Master: Gilles Dowek has given a course at MPRI.
- Master: Gilles Dowek is in charge of the second year of Masters at the École normale de Paris Saclay.
- Master: Gilles Dowek has given a one week invited course at the University of Buenos Aires.
- Licence: Guillaume Genestier, logic tutorials, 45h, L3, ENS Paris-Saclay
- Licence: Guillaume Genestier, complexity remedial classes, 11h, L3, ENS Paris-Saclay
- Licence: Gaspard Férey, language theory, 44h, L3, EISTI
- Licence: François Thiré, (spring) logic project, 26h, L3 ENS Paris-Saclay
- Licence: François Thiré, (spring) Programmation 2 tutorials, 26h, L3 ENS Paris-Saclay
- Licence: François Thiré, (fall) Architectures and Systems tutorials, 36h, L3 ENS Paris-Saclay
- Frédéric Blanqui is co-director of the pole 4 of the doctoral school STIC of the University Paris-Saclay.
- Frédéric Blanqui is member of the committee of the doctoral school of the ENS Paris-Saclay.
- Frédéric Blanqui is in charge of following PhD students at LSV.

9.2.2. Supervision

- PhD Defended: Frédéric Gilbert, Gilles Dowek and Florent Kirchner,
- PhD in progress: Guillaume Bury, David Delahaye and Gilles Dowek,
- PhD in progress: Guillaume Genestier, termination in $\lambda\Pi$-calculus modulo theory, 01/10/17, Frédéric Blanqui and Olivier Hermant,
- PhD in progress: Mohamed Yacine El Haddad, using automated provers in proof assistants, 05/01/18, Frédéric Blanqui and Guillaume Burel,

- PhD in progress: Gaspard Férey, Associative-Commutative rewriting in the $\lambda\Pi$-calculus, 01/09/18, Gilles Dowek,
- PhD in progress: François Thiré, Design tools to make interoperability easier in DEDUKTI, 01/09/18, Gilles Dowek.

9.2.3. Juries

Gilles Dowek has been a member of the Jury of the PhD defence of Pierre Boutry. He has been an evaluator of the thesis of Thibault Gauthier. He has been a member of the Jury of the habilitation defence of Julien Signoles and of Alexei Grinbaum.

9.3. Popularization

9.3.1. Articles and contents

Gilles Dowek writes a monthly column in Pour la Science (12 issues) and has started a bi-monthly column in Le Monde (3 issues).

Gilles Dowek has given interviews to France Inter, Radio France Internationale, France Culture, Ouest France, Usbek et Rica, and Philosophie Magazine.

9.3.2. Education

Gilles Dowek has participated to meetings on scientific education in Switzerland, Belgium, and Côte d'Ivoire.

He has been heard by a committee of the the Éducation Nationale on pedagogical data and privacy.

He has given a talk on job mutations to mathematics inspectors.

9.3.3. Interventions

Gilles Dowek has given popular science talks in Toulouse, Antony, Issoudun, Rueil Malmaison, Saint Louis, Saint-Cloud, Rennes, Nancy, Paris, Nîmes, St Quentin en Yvelines, Montbéliard, Molaix, St Agrève, Rhodes, Marcoule, and Juvisy.

10. Bibliography

Major publications by the team in recent years

[1] A. ASSAF, G. BUREL, R. CAUDERLIER, D. DELAHAYE, G. DOWEK, C. DUBOIS, F. GILBERT, P. HALMAGRAND, O. HERMANT, R. SAILLARD.*Expressing theories in the $\lambda\Pi$-calculus modulo theory and in the Dedukti system*, in "22nd International Conference on Types for Proofs and Programs, TYPES 2016", Novi SAd, Serbia, May 2016, https://hal-mines-paristech.archives-ouvertes.fr/hal-01441751

[2] B. BARRAS, T. COQUAND, S. HUBER.*A generalization of the Takeuti-Gandy interpretation*, in "Mathematical Structures in Computer Science", 2015, vol. 25, n^{o} 5, p. 1071–1099, https://doi.org/10.1017/S0960129514000504

[3] F. BLANQUI.*Definitions by rewriting in the Calculus of Constructions*, in "Mathematical Structures in Computer Science", 2005, vol. 15, n^{o} 1, p. 37-92 [*DOI :* 10.1017/S0960129504004426], http://hal.inria.fr/inria-00105648/en/

[4] F. BLANQUI, J.-P. JOUANNAUD, A. RUBIO.*The Computability Path Ordering*, in "Logical Methods in Computer Science", October 2015 [*DOI :* 10.2168/LMCS-11(4:3)2015], https://hal.inria.fr/hal-01163091

[5] G. BUREL.*Experimenting with Deduction Modulo*, in "CADE 2011", V. SOFRONIE-STOKKERMANS, N. BJØRNER (editors), Lecture Notes in Artificial Intelligence, Springer, 2011, vol. 6803, p. 162–176

[6] D. COUSINEAU, G. DOWEK.*Embedding Pure Type Systems in the $\lambda\Pi$-calculus modulo*, in "Typed lambda calculi and applications", S. RONCHI DELLA ROCCA (editor), Lecture Notes in Computer Science, Springer-Verlag, 2007, vol. 4583, p. 102-117

[7] G. DOWEK, T. HARDIN, C. KIRCHNER.*Theorem proving modulo*, in "Journal of Automated Reasoning", 2003, vol. 31, p. 33-73

[8] O. HERMANT.*Resolution is Cut-Free*, in "Journal of Automated Reasoning", March 2010, vol. 44, n^{o} 3, p. 245-276

[9] M. JACQUEL, K. BERKANI, D. DELAHAYE, C. DUBOIS.*Verifying B Proof Rules using Deep Embedding and Automated Theorem Proving*, in "Software and Systems Modeling (SoSyM)", June 2013

[10] M. JACQUEL, K. BERKANI, D. DELAHAYE, C. DUBOIS.*Tableaux Modulo Theories Using Superdeduction*, in "Global Journal of Advanced Software Engineering (GJASE)", December 2014, vol. 1, p. 1 - 13 [*DOI :* 10.1007/978-3-642-31365-3_26], https://hal.archives-ouvertes.fr/hal-01099338

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] F. GILBERT.*Extending higher-order logic with predicate subtyping: Application to PVS*, Université Sorbonne Paris Cité ; Université Paris Diderot, April 2018, https://hal.inria.fr/hal-01673518

Articles in International Peer-Reviewed Journal

[12] F. BLANQUI.*Size-based termination of higher-order rewriting*, in "Journal of Functional Programming", April 2018, https://arxiv.org/abs/1802.06603 [*DOI :* 10.1017/S0956796818000072], https://hal.inria.fr/hal-01424921

[13] Y. JIANG, J. LIU, G. DOWEK, K. JI.*Towards Combining Model Checking and Proof Checking*, in "The Computer Journal", 2019, https://hal.inria.fr/hal-01970274

[14] R. LEPIGRE, C. RAFFALLI.*Abstract Representation of Binders in OCaml using the Bindlib Library*, in "Electronic Proceedings in Theoretical Computer Science", July 2018, vol. 274, p. 42-56, https://arxiv.org/abs/1807.01872 - In Proceedings LFMTP 2018, arXiv:1807.01352 [*DOI :* 10.4204/EPTCS.274.4], https://hal.inria.fr/hal-01972050

[15] F. THIRÉ.*Sharing a Library between Proof Assistants: Reaching out to the HOL Family* *, in "Electronic Proceedings in Theoretical Computer Science", July 2018, vol. 274, p. 57 - 71 [*DOI :* 10.4204/EPTCS.274.5], https://hal.inria.fr/hal-01929714

International Conferences with Proceedings

[16] F. BLANQUI, G. GENESTIER.*Termination of $\lambda\Pi$ modulo rewriting using the size-change principle (work in progress)*, in "16th International Workshop on Termination", Oxford, United Kingdom, S. LUCAS (editor), July 2018, p. 10-14, https://arxiv.org/abs/1812.01853 , https://hal.inria.fr/hal-01944731

[17] G. BUREL.*Linking Focusing and Resolution with Selection*, in "43rd International Symposium on Mathematical Foundations of Computer Science (MFCS 2018)", Liverpool, United Kingdom, I. POTAPOV, P. SPIRAKIS, J. WORRELL (editors), Leibniz International Proceedings in Informatics (LIPIcs), Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, August 2018, vol. 117, p. 9:1-9:14 [*DOI :* 10.4230/LIPICS.MFCS.2018.9], https://hal.inria.fr/hal-01670476

[18] N. DERSHOWITZ, J.-P. JOUANNAUD.*Drags: A Simple Algebraic Framework For Graph Rewriting*, in "TERMGRAPH 2018 - 10th International Workshop on Computing with Terms and Graphs", Oxford, United Kingdom, July 2018, https://hal.inria.fr/hal-01853836

National Conferences with Proceeding

[19] S. COLIN, R. LEPIGRE, G. SCHERER.*Unboxing Mutually Recursive Type Definitions in OCaml*, in "JFLA 2019", Les Rousses, France, January 2019, https://arxiv.org/abs/1811.02300 , https://hal.inria.fr/hal-01929508

Conferences without Proceedings

[20] N. DERSHOWITZ, J.-P. JOUANNAUD.*Graph Path Orderings*, in "2nd International Conference on Logic for Programming, Artificial Intelligence and Reasoning", Awassa, Ethiopia, G. BARTHE, G. SUTCLIFFE, M. VEANES (editors), EPiC Series in Computing, November 2018, vol. 57, p. 1-18, https://hal.inria.fr/hal-01903086

Research Reports

[21] Q. YE.*Comparaison des termes avec partage*, LSV, ENS Cachan, CNRS, Inria, Université Paris-Saclay, Cachan (France), July 2018, https://hal.inria.fr/hal-01973539

Other Publications

[22] F. BLANQUI, G. GENESTIER, O. HERMANT.*Dependency Pairs Termination in Dependent Type Theory Modulo Rewriting*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01943941

[23] A. DÍAZ-CARO, G. DOWEK.*A logic identifying isomorphic propositions*, August 2018, https://arxiv.org/abs/1501.06125 - There is a mistake in the main proof (strong normalisation), and the system is actually not normalising in its current form, https://hal.inria.fr/hal-01109104

[24] F. GILBERT.*Verifiable certificates for predicate subtyping*, January 2019, working paper or preprint, https://hal.inria.fr/hal-01977585

[25] I. LACHHEB.*Une interface pour Dedukti*, LSV, ENS Cachan, CNRS, Inria, Université Paris-Saclay, Cachan (France), September 2018, https://hal.inria.fr/hal-01898401

[26] W. MOUSTAOUI.*Encyclopédie en ligne de démonstrations formelles*, LSV, ENS Cachan, CNRS, Inria, Université Paris-Saclay, Cachan (France), September 2018, https://hal.inria.fr/hal-01975446

[27] A.-D. PAPADOPOULOS.*Industrial Placement Report Research Placement in Deducteam*, Imperial College London, October 2018, https://hal.inria.fr/hal-01890253

References in notes

[28] A. ASSAF.*A framework for defining computational higher-order logics*, École polytechnique, September 2015, https://pastel.archives-ouvertes.fr/tel-01235303

[29] M. BOESPFLUG.*Conception d'un noyau de vérification de preuves pour le λΠ-calcul modulo*, École Polytechnique, 2011

[30] G. BURY, S. CRUANES, D. DELAHAYE, P. EUVRARD.*An Automation-Friendly Set Theory for the B Method*, in "Abstract State Machines, Alloy, B, TLA, VDM, and Z - 6th International Conference, ABZ 2018, Southampton, UK, June 5-8, 2018, Proceedings", M. J. BUTLER, A. RASCHKE, T. S. HOANG, K. REICHL (editors), Lecture Notes in Computer Science, Springer, 2018, vol. 10817, p. 409–414, https://doi.org/10.1007/978-3-319-91271-4_32

[31] G. DOWEK.*A Theory Independent Curry-de Bruijn-howard Correspondence*, in "Proceedings of the 39th International Colloquium Conference on Automata, Languages, and Programming - Volume Part II", Berlin, Heidelberg, ICALP'12, Springer-Verlag, 2012, p. 13–15, http://dx.doi.org/10.1007/978-3-642-31585-5_2

[32] R. LEPIGRE.*PML 2 : Integrated Program Verification in ML*, in "23rd International Conference on Types for Proofs and Programs (TYPES 2017)", Budapest, Hungary, July 2017, 27 [*DOI :* 10.4230/LIPIcs.TYPES.2017.5], https://hal.inria.fr/hal-01972000

Project-Team DEFI

Shape reconstruction and identification

IN COLLABORATION WITH: Centre de Mathématiques Appliquées (CMAP)

IN PARTNERSHIP WITH:
Ecole Polytechnique

RESEARCH CENTER
Saclay - Île-de-France

THEME
Numerical schemes and simulations

Table of contents

Project-Team DEFI

Creation of the Project-Team: 2009 January 01

Keywords:

Computer Science and Digital Science:

A6. - Modeling, simulation and control
A6.1. - Methods in mathematical modeling
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.4. - Multiscale modeling
A6.2. - Scientific computing, Numerical Analysis & Optimization
A6.2.1. - Numerical analysis of PDE and ODE
A6.2.3. - Probabilistic methods
A6.2.6. - Optimization
A6.2.7. - High performance computing
A6.3. - Computation-data interaction
A6.3.1. - Inverse problems
A6.3.5. - Uncertainty Quantification

Other Research Topics and Application Domains:

B1.2.1. - Understanding and simulation of the brain and the nervous system
B1.2.3. - Computational neurosciences
B2.6.1. - Brain imaging
B3.2. - Climate and meteorology
B3.3.4. - Atmosphere

1. Team, Visitors, External Collaborators

Research Scientists

Houssem Haddar [Team leader, Inria, Senior Researcher]
Marcella Bonazzoli [Inria, Researcher, from Oct 2018]
Lucas Chesnel [Inria, Researcher]
Pietro-Marco Congedo [Inria, Researcher, from Feb 2018, HDR]
Jing-Rebecca Li [Inria, Researcher, HDR]

Faculty Member

Grégoire Allaire [Ecole polytechnique, Professor, HDR]

External Collaborators

Lorenzo Audibert [EDF]
Olivier Le Maitre [CNRS, from Apr 2018]

PhD Students

Julien Carlier [Ecole polytechnique, from Apr 2018]
Joao Miguel Felicio Dos Reis [Inria, from Apr 2018]
Hugo Girardon [EDF]
Giulio Gori [Inria, from Apr 2018]
Jingjing Hao [Inria]
Kevish Napal [Ecole polytechnique]

Mahran Rihani [Ecole polytechnique]
Marwa Kchaou [ENIT]

Post-Doctoral Fellows

Marc Bakry [Inria, until Oct 2018]
Xiaoli Liu [Inria, from Dec 2018]
Imen Mekkaoui [Inria]
Thi Phong Nguyen [Inria, until Feb 2018]

Visiting Scientists

Barbara Arizmendi [Inria, from Sep 2018]
Fioralba Cakoni [Rutgers University, Jul 2018]
Ana Isabel Del Val Benitez [PhD, Ecole Polytechnique and von Karman Institute, from Apr 2018 until May 2018]
Christian Sabater Campomanes [DLR, Germany, visiting PhD student, from Sep 2018 until Nov 2018]

Administrative Assistant

Natalia Alves [Inria]

2. Overall Objectives

2.1. Overall Objectives

The research activity of our team is dedicated to the design, analysis and implementation of efficient numerical methods to solve inverse and shape/topological optimization problems, eventually including system uncertainties, in connection with acoustics, electromagnetism, elastodynamics, diffusion, and fluid mechanics.

Sought practical applications include radar and sonar applications, bio-medical imaging techniques, non-destructive testing, structural design, composite materials, diffusion magnetic resonance imaging, fluid-driven applications in aerospace/energy fields.

Roughly speaking, the model problem consists in determining information on, or optimizing the geometry (topology) and the physical properties of unknown targets from given constraints or measurements, for instance, measurements of diffracted waves or induced magnetic fields. Moreover, system uncertainties can ben systematically taken into account to provide a measure of confidence of the numerical prediction.

In general this kind of problems is non-linear. The inverse ones are also severely ill-posed and therefore require special attention from regularization point of view, and non-trivial adaptations of classical optimization methods.

Our scientific research interests are the following:

- Theoretical understanding and analysis of the forward and inverse mathematical models, including in particular the development of simplified models for adequate asymptotic configurations.
- The design of efficient numerical optimization/inversion methods which are quick and robust with respect to noise. Special attention will be paid to algorithms capable of treating large scale problems (e.g. 3-D problems) and/or suited for real-time imaging.
- Propose new methods and develop advanced tools to perform uncertainty quantification for optimization/inversion.
- Development of prototype softwares for specific applications or tutorial toolboxes.

We were particularly interested in the development of the following themes

- Qualitative and quantitative methods for inverse scattering problems
- Topological optimization methods
- Forward and inverse models for Diffusion MRI
- Forward/Backward uncertainty quantification methods for optimization/inversion problems in the context of expensive computer codes.

3. Research Program

3.1. Research Program

The research activity of our team is dedicated to the design, analysis and implementation of efficient numerical methods to solve inverse and shape/topological optimization problems, eventually including system uncertainties, in connection with wave imaging, structural design, non-destructive testing and medical imaging modalities. We are particularly interested in the development of fast methods that are suited for real-time applications and/or large scale problems. These goals require to work on both the physical and the mathematical models involved and indeed a solid expertise in related numerical algorithms. A part of the research activity is also devoted to take into account system uncertainties in the solving of inverse/optimization problems. At the interface of physics, mathematics, and computer science, Uncertainty Quantification (UQ) focuses on the development of frameworks and methods to characterize uncertainties in predictive computations. Uncertainties and errors arise at different stages of the numerical simulation. First, errors are introduced due to the physical simplifications in the mathematical modeling of the system investigated; other errors come from the numerical resolution of the mathematical model, due in particular to finite discretization and computations with finite accuracy and tolerance; finally, errors are due a limited knowledge of input quantities (parameters) appearing in the definition of the numerical model being solved.

This section intends to give a general overview of our research interests and themes. We choose to present them through the specific academic example of inverse scattering problems (from inhomogeneities), which is representative of foreseen developments on both inversion and (topological) optimization methods. The practical problem would be to identify an inclusion from measurements of diffracted waves that result from the interaction of the sought inclusion with some (incident) waves sent into the probed medium. Typical applications include biomedical imaging where using micro-waves one would like to probe the presence of pathological cells, or imaging of urban infrastructures where using ground penetrating radars (GPR) one is interested in finding the location of buried facilities such as pipelines or waste deposits. This kind of applications requires in particular fast and reliable algorithms.

By “imaging” we refer to the inverse problem where the concern is only the location and the shape of the inclusion, while “identification” may also indicate getting informations on the inclusion physical parameters.

Both problems (imaging and identification) are non linear and ill-posed (lack of stability with respect to measurements errors if some careful constrains are not added). Moreover, the unique determination of the geometry or the coefficients is not guaranteed in general if sufficient measurements are not available. As an example, in the case of anisotropic inclusions, one can show that an appropriate set of data uniquely determine the geometry but not the material properties.

These theoretical considerations (uniqueness, stability) are not only important in understanding the mathematical properties of the inverse problem, but also guide the choice of appropriate numerical strategies (which information can be stably reconstructed) and also the design of appropriate regularization techniques. Moreover, uniqueness proofs are in general constructive proofs, i.e. they implicitly contain a numerical algorithm to solve the inverse problem, hence their importance for practical applications. The sampling methods introduced below are one example of such algorithms.

A large part of our research activity is dedicated to numerical methods applied to the first type of inverse problems, where only the geometrical information is sought. In its general setting the inverse problem is very challenging and no method can provide universally satisfying solution (respecting the balance cost-precision-stability). This is why in the majority of the practically employed algorithms, some simplification of the underlying mathematical model is used, according to the specific configuration of the imaging experiment. The most popular ones are geometric optics (the Kirchhoff approximation) for high frequencies and weak scattering (the Born approximation) for small contrasts or small obstacles. They actually give full satisfaction for a wide range of applications as attested by the large success of existing imaging devices (radar, sonar, ultrasound, X-ray tomography, etc.), that rely on one of these approximations.

In most cases, the used simplification result in a linearization of the inverse problem and therefore is usually valid only if the latter is weakly non-linear. The development of simplified models and the improvement of their efficiency is still a very active research area. With that perspective, we are particularly interested in deriving and studying higher order asymptotic models associated with small geometrical parameters such as: small obstacles, thin coatings, wires, periodic media, ... Higher order models usually introduce some non linearity in the inverse problem, but are in principle easier to handle from the numerical point of view than in the case of the exact model.

A larger part of our research activity is dedicated to algorithms that avoid the use of such approximations and that are efficient where classical approaches fail: i.e. roughly speaking when the non linearity of the inverse problem is sufficiently strong. This type of configuration is motivated by the applications mentioned below, and occurs as soon as the geometry of the unknown media generates non negligible multiple scattering effects (multiply-connected and closely spaces obstacles) or when the used frequency is in the so-called resonant region (wave-length comparable to the size of the sought medium). It is therefore much more difficult to deal with and requires new approaches. Our ideas to tackle this problem is mainly motivated and inspired by recent advances in shape and topological optimization methods and in so-called sampling methods.

Sampling methods are fast imaging solvers adapted to multi-static data (multiple receiver-transmitter pairs) at a fixed frequency. Even if they do not use any linearization the forward model, they rely on computing the solutions to a set of linear problems of small size, that can be performed in a completely parallel procedure. Our team has already a solid expertise in these methods applied to electromagnetic 3-D problems. The success of such approaches was their ability to provide a relatively quick algorithm for solving 3-D problems without any need for a priori knowledge on the physical parameters of the targets. These algorithms solve only the imaging problem, in the sense that only the geometrical information is provided.

Despite the large efforts already spent in the development of this type of methods, either from the algorithmic point of view or the theoretical one, numerous questions are still open. These attractive new algorithms also suffer from the lack of experimental validations, due to their relatively recent introduction. We also would like to invest on this side by developing collaborations with engineering research groups that have experimental facilities. From the practical point of view, the most potential limitation of sampling methods would be the need of a large amount of data to achieve a reasonable accuracy. On the other hand, optimization methods do not suffer from this constrain but they require good initial guess to ensure convergence and reduce the number of iterations. Therefore it seems natural to try to combine the two class of methods in order to calibrate the balance between cost and precision.

Among various shape optimization methods, the Level Set method seems to be particularly suited for such a coupling. First, because it shares similar mechanism as sampling methods: the geometry is captured as a level set of an “indicator function” computed on a cartesian grid. Second, because the two methods do not require any a priori knowledge on the topology of the sought geometry. Beyond the choice of a particular method, the main question would be to define in which way the coupling can be achieved. Obvious strategies consist in using one method to pre-process (initialization) or post-process (find the level set) the other. But one can also think of more elaborate ones, where for instance a sampling method can be used to optimize the choice of the incident wave at each iteration step. The latter point is closely related to the design of so called “focusing incident waves” (which are for instance the basis of applications of the time-reversal principle). In the frequency regime, these incident waves can be constructed from the eigenvalue decomposition of the data operator used by sampling methods. The theoretical and numerical investigations of these aspects are still not completely understood for electromagnetic or elastodynamic problems.

Other topological optimization methods, like the homogenization method or the topological gradient method, can also be used, each one provides particular advantages in specific configurations. It is evident that the development of these methods is very suited to inverse problems and provide substantial advantage compared to classical shape optimization methods based on boundary variation. Their applications to inverse problems has not been fully investigated. The efficiency of these optimization methods can also be increased for adequate asymptotic configurations. For instance small amplitude homogenization method can be used as an efficient relaxation method for the inverse problem in the presence of small contrasts. On the other hand, the topological gradient method has shown to perform well in localizing small inclusions with only one iteration.

A broader perspective would be the extension of the above mentioned techniques to time-dependent cases. Taking into account data in time domain is important for many practical applications, such as imaging in cluttered media, the design of absorbing coatings or also crash worthiness in the case of structural design.

For the identification problem, one would like to also have information on the physical properties of the targets. Of course optimization methods is a tool of choice for these problems. However, in some applications only a qualitative information is needed and obtaining it in a cheaper way can be performed using asymptotic theories combined with sampling methods. We also refer here to the use of so called transmission eigenvalues as qualitative indicators for non destructive testing of dielectrics.

We are also interested in parameter identification problems arising in diffusion-type problems. Our research here is mostly motivated by applications to the imaging of biological tissues with the technique of Diffusion Magnetic Resonance Imaging (DMRI). Roughly speaking DMRI gives a measure of the average distance travelled by water molecules in a certain medium and can give useful information on cellular structure and structural change when the medium is biological tissue. In particular, we would like to infer from DMRI measurements changes in the cellular volume fraction occurring upon various physiological or pathological conditions as well as the average cell size in the case of tumor imaging. The main challenges here are 1) correctly model measured signals using diffusive-type time-dependent PDEs 2) numerically handle the complexity of the tissues 3) use the first two to identify physically relevant parameters from measurements. For the last point we are particularly interested in constructing reduced models of the multiple-compartment Bloch-Torrey partial differential equation using homogenization methods.

The Team devotes a large effort focused on the formulation, implementation and validation of numerical methods for using scientific computing to drive experiments and available data (coming from models, simulation and experiments) by taking into account the system uncertainty. The team is also invested in exploiting the intimate relationship between optimisation and UQ to make Optimisation Under Uncertainty (OUU) tractable. A part of these activities is declined to the simulation of high-fidelity models for fluids, in three main fields, aerospace, energy and environment.

The Team is working on developing original UQ representations and algorithms to deal with complex and large scale models, having high dimensional input parameters with complexes influences. We are organizing our core research activities along different methodological UQ developments related to the challenges discussed above. Obviously, some efforts are shared by different initiatives or projects, and some of them include the continuous improvement of the non-intrusive methods constituting our software libraries. These actions are not detailed in the following, to focus the presentation on more innovative aspects, but we mentioned nonetheless the continuous developments and incorporation in our libraries of advanced sparse grid methods, sparsity promoting strategies and low rank methods.

An effort is dedicated to the efficient construction of surrogate models that are central in both forward and backward UQ problems, aiming at large-scale simulations relevant to engineering applications, with high dimensional input parameters.

Sensitivity analyses and other forward UQ problems (e.g., estimation of failure probabilities, rare events,. . .) depends on the input uncertainty model. Most often, for convenience or because of the lack of data, the independence of the uncertain inputs is assumed. In the Team, we are investigating approaches dedicated to a) the construction of uncertainty models that integrate the available information and expert knowledge(s) in a consistent and objective fashion. To this end, several mathematical frameworks are already available, e.g the maximum entropy principle, likelihood maximization and moment matching methods, but their application to real engineering problems remains scarce and their systematic use raises multiple challenges, both to construct the uncertainty model and to solve the related UQ problems (forward and backward). Because of the importance of the available data and expertise to build the model, the contributions of the Team in these areas depend on the needs and demands of end-users and industrial partners.

To mitigate computational complexity, the Team is exploring multi-fidelity approaches in the context of expensive simulations. We combine predictions of models with different levels of discretizations and physical simplifications to construct, at a controlled cost, reliable surrogate models of simulation outputs or directly objective functions and possibly constraints, to enable the resolution of robust optimization and stochastic inverse problems. Again, one difficulty to be addressed by the Team is the design of the computer experiments to obtain the best multi-fidelity model at the lowest cost (of for a prescribed computational budgets), with respect to the end use of the model. The last point is particularly challenging as it calls for accuracy for output values that are usually unknown a priori but must be estimated as the model construction proceeds.

4. Application Domains

4.1. Radar and GPR applications

Conventional radar imaging techniques (ISAR, GPR, etc.) use backscattering data to image targets. The commonly used inversion algorithms are mainly based on the use of weak scattering approximations such as the Born or Kirchhoff approximation leading to very simple linear models, but at the expense of ignoring multiple scattering and polarization effects. The success of such an approach is evident in the wide use of synthetic aperture radar techniques.

However, the use of backscattering data makes 3-D imaging a very challenging problem (it is not even well understood theoretically) and as pointed out by Brett Borden in the context of airborne radar: “In recent years it has become quite apparent that the problems associated with radar target identification efforts will not vanish with the development of more sensitive radar receivers or increased signal-to-noise levels. In addition it has (slowly) been realized that greater amounts of data - or even additional “kinds” of radar data, such as added polarization or greatly extended bandwidth - will all suffer from the same basic limitations affiliated with incorrect model assumptions. Moreover, in the face of these problems it is important to ask how (and if) the complications associated with radar based automatic target recognition can be surmounted.” This comment also applies to the more complex GPR problem.

Our research themes will incorporate the development, analysis and testing of several novel methods, such as sampling methods, level set methods or topological gradient methods, for ground penetrating radar application (imaging of urban infrastructures, landmines detection, underground waste deposits monitoring,) using multistatic data.

4.2. Biomedical imaging

Among emerging medical imaging techniques we are particularly interested in those using low to moderate frequency regimes. These include Microwave Tomography, Electrical Impedance Tomography and also the closely related Optical Tomography technique. They all have the advantage of being potentially safe and relatively cheap modalities and can also be used in complementarity with well established techniques such as X-ray computed tomography or Magnetic Resonance Imaging.

With these modalities tissues are differentiated and, consequentially can be imaged, based on differences in dielectric properties (some recent studies have proved that dielectric properties of biological tissues can be a strong indicator of the tissues functional and pathological conditions, for instance, tissue blood content, ischemia, infarction, hypoxia, malignancies, edema and others). The main challenge for these functionalities is to built a 3-D imaging algorithm capable of treating multi-static measurements to provide real-time images with highest (reasonably) expected resolutions and in a sufficiently robust way.

Another important biomedical application is brain imaging. We are for instance interested in the use of EEG and MEG techniques as complementary tools to MRI. They are applied for instance to localize epileptic centers or active zones (functional imaging). Here the problem is different and consists into performing passive imaging: the epileptic centers act as electrical sources and imaging is performed from measurements of induced currents. Incorporating the structure of the skull is primordial in improving the resolution of the imaging procedure. Doing this in a reasonably quick manner is still an active research area, and the use of asymptotic models would offer a promising solution to fix this issue.

4.3. Non destructive testing and parameter identification

One challenging problem in this vast area is the identification and imaging of defaults in anisotropic media. For instance this problem is of great importance in aeronautic constructions due to the growing use of composite materials. It also arises in applications linked with the evaluation of wood quality, like locating knots in timber in order to optimize timber-cutting in sawmills, or evaluating wood integrity before cutting trees. The anisotropy of the propagative media renders the analysis of diffracted waves more complex since one cannot only relies on the use of backscattered waves. Another difficulty comes from the fact that the micro-structure of the media is generally not well known a priori.

Our concern will be focused on the determination of qualitative information on the size of defaults and their physical properties rather than a complete imaging which for anisotropic media is in general impossible. For instance, in the case of homogeneous background, one can link the size of the inclusion and the index of refraction to the first eigenvalue of so-called interior transmission problem. These eigenvalues can be determined form the measured data and a rough localization of the default. Our goal is to extend this kind of idea to the cases where both the propagative media and the inclusion are anisotropic. The generalization to the case of cracks or screens has also to be investigated.

In the context of nuclear waste management many studies are conducted on the possibility of storing waste in a deep geological clay layer. To assess the reliability of such a storage without leakage it is necessary to have a precise knowledge of the porous media parameters (porosity, tortuosity, permeability, etc.). The large range of space and time scales involved in this process requires a high degree of precision as well as tight bounds on the uncertainties. Many physical experiments are conducted in situ which are designed for providing data for parameters identification. For example, the determination of the damaged zone (caused by excavation) around the repository area is of paramount importance since microcracks yield drastic changes in the permeability. Level set methods are a tool of choice for characterizing this damaged zone.

4.4. Diffusion MRI

In biological tissues, water is abundant and magnetic resonance imaging (MRI) exploits the magnetic property of the nucleus of the water proton. The imaging contrast (the variations in the grayscale in an image) in standard MRI can be from either proton density, T1 (spin-lattice) relaxation, or T2 (spin-spin) relaxation and the contrast in the image gives some information on the physiological properties of the biological tissue at different physical locations of the sample. The resolution of MRI is on the order of millimeters: the greyscale value shown in the imaging pixel represents the volume-averaged value taken over all the physical locations contained that pixel.

In diffusion MRI, the image contrast comes from a measure of the average distance the water molecules have moved (diffused) during a certain amount of time. The Pulsed Gradient Spin Echo (PGSE) sequence is a commonly used sequence of applied magnetic fields to encode the diffusion of water protons. The term 'pulsed' means that the magnetic fields are short in duration, an the term gradient means that the magnetic fields vary linearly in space along a particular direction. First, the water protons in tissue are labelled with nuclear spin at a precession frequency that varies as a function of the physical positions of the water molecules via the application of a pulsed (short in duration, lasting on the order of ten milliseconds) magnetic field. Because the precessing frequencies of the water molecules vary, the signal, which measures the aggregate phase of the water molecules, will be reduced due to phase cancellations. Some time (usually tens of milliseconds) after the first pulsed magnetic field, another pulsed magnetic field is applied to reverse the spins of the water molecules. The time between the applications of two pulsed magnetic fields is called the 'diffusion time'. If the water molecules have not moved during the diffusion time, the phase dispersion will be reversed, hence the signal loss will also be reversed, the signal is called refocused. However, if the molecules have moved during the diffusion time, the refocusing will be incomplete and the signal detected by the MRI scanner if weaker than if the water molecules have not moved. This lack of complete refocusing is called the signal attenuation and is the basis of the image contrast in DMRI. the pixels showning more signal attenuation is associated with further water displacement during the diffusion time, which may be linked to physiological factors, such as higher cell membrane permeability, larger cell sizes, higher extra-cellular volume fraction.

We model the nuclear magnetization of water protons in a sample due to diffusion-encoding magnetic fields by a multiple compartment Bloch-Torrey partial differential equation, which is a diffusive-type time-dependent PDE. The DMRI signal is the integral of the solution of the Bloch-Torrey PDE. In a homogeneous medium, the intrinsic diffusion coeffcient D will appear as the slope of the semi-log plot of the signal (in approporiate units). However, because during typical scanning times, 50-100ms, water molecules have had time to travel a diffusion distance which is long compared to the average size of the cells, the slope of the semi-log plot of the signal is in fact a measure of an 'effective' diffusion coefficient. In DMRI applications, this measured quantity is called the 'apparent diffusion coefficient' (ADC) and provides the most commonly used form the image contrast for DMRI. This ADC is closely related to the effective diffusion coefficient obtainable from mathematical homogenization theory.

4.5. Fluid flow applications

Specific actions are devoted to the problem of atmospheric reentry simulations. We focus on several aspects : i) on the development of innovative algorithms improving the prediction of hypersonic flows and including system uncertainties, ii) on the application of these methods to the atmospheric reentry of space vehicles for the control and the optimization of the trajectory, iii) on the debris reentry, which is of fundamental importance for NASA, CNES and ESA. Several works are already initiated with funding from CNES, Thales, and ASL. An ongoing activity concerns the design of the Thermal Protection System (TPS) that shields the spacecraft from aerothermal heating, generated by friction at the surface of the vehicle. The TPS is usually composed of different classes of materials, depending on the mission and the planned trajectory. One major issue is to model accurately the material response to ensure a safe design. High-fidelity material modeling for ablative materials has been developed by NASA, but a lot of work is still needed concerning the assessment of physical and modeling uncertainties during the design process. Our objective is to set up a predictive numerical tool to reliably estimate the response of ablative materials for different aerothermal conditions.

An important effort is dedicated to the simulation of fluids featuring complex thermodynamic behavior, in the context of two distinct projects: the VIPER project, funded by Aquitaine Region, and a project with CWI (Scientific Computing Group). Dense gases (DGs) are defined as single-phase vapors operating at temperatures and pressures conditions close to the saturation curve. The interest in studying complex dynamics of compressible dense gas flows comes from the potential technological advantages of using these fluids in energy conversion cycles, such as in Organic Rankine Cycles (ORCs) which used dense gases as energy converters for biomass fuels and low-grade heat from geothermal or industrial waste heat sources. Since these fluids feature large uncertainties in their estimated thermodynamic properties (critical properties, acentric factor, etc.), a meaningful numerical prediction of the performance must necessarily take into account these uncertainties. Other sources of uncertainties include, but are not limited to, the inlet boundary conditions which are often unknown in dense gases applications. Moreover, a robust optimization must also include the more generic uncertainty introduced by the machining tolerance in the construction of the turbine blades.

5. Highlights of the Year

5.1. Highlights of the Year

- Fellowship for the participation to the Center for Turbulence Research Summer Program in Stanford University, June-July 2018 (PM Congedo, G. Gori).

6. New Software and Platforms

6.1. samplings-2d

This software solves forward and inverse problems for the Helmholtz equation in 2-D.

FUNCTIONAL DESCRIPTION: This software is written in Fortran 90 and is related to forward and inverse problems for the Helmholtz equation in 2-D. It includes three independent components. * The first one solves to scattering problem using integral equation approach and supports piecewise-constant dielectrics and obstacles with impedance boundary conditions. * The second one contains various samplings methods to solve the inverse scattering problem (LSM, RGLSM(s), Factorization, MuSiC) for near-field or far-field setting. * The third component is a set of post processing functionalities to visualize the results

- Participant: Houssem Haddar
- Contact: Houssem Haddar
- URL: http://sourceforge.net/projects/samplings-2d/

6.2. SAXS-LMA-HSPY

SAXS inversion using LMA and HSPY models
KEYWORD: SAXS measurements
FUNCTIONAL DESCRIPTION: This software determines nanoparticles size distribution from SAXS measurements (Small Angle X-ray Scattering). It contains two different approaches. The first one is based on a linear LMA model with automatic search for model parameters. The second approach uses a non-linear inversion of the HSPY model.

- Authors: Marc Bakry and Houssem Haddar
- Contact: Marc Bakry

6.3. FVforBlochTorrey

KEYWORDS: Simulation - PDE - Diffusion imaging - MRI
FUNCTIONAL DESCRIPTION: We developed a Matlab toolbox for solving the multiple-compartments Bloch-Torrey partial differential equation in 3D to simulate the water proton magnetization of a sample under the influence of diffusion-encoing magnetic field gradient pulses. We coupled the finite element spatial discretization with several ODE solvers in time that are available inside Matlab.

Result: the code will be made available on GitHub in 2019.

- Participant: Jing Rebecca Li
- Contact: Jing Rebecca Li

7. New Results

7.1. Qualitative and quantitative methods for inverse problems

7.1.1. On the Factorization Method for a Far Field Inverse Scattering Problem in the Time Domain

F. Cakoni, H. Haddar and A. Lechleiter

We develop a factorization method to obtain explicit characterization of a (possibly non-convex) Dirichlet scattering object from measurements of time-dependent causal scattered waves in the far field regime. In particular, we prove that far fields of solutions to the wave equation due to particularly modified incident waves, characterize the obstacle by a range criterion involving the square root of the time derivative of the corresponding far field operator. Our analysis makes essential use of a coercivity property of the solution of the Dirichlet initial boundary value problem for the wave equation in the Laplace domain. This forces us to consider this particular modification of the far field operator. The latter in fact, can be chosen arbitrarily close to the true far field operator given in terms of physical measurements.

7.1.2. New interior transmission problem applied to a single Floquet–Bloch mode imaging of local perturbations in periodic media

F. Cakoni, H. Haddar and T.P Nguyen

We consider the imaging of local perturbations of an infinite penetrable periodic layer. A cell of this periodic layer consists of several bounded inhomogeneities situated in a known homogeneous media. We use a differential linear sampling method to reconstruct the support of perturbations without using the Green's function of the periodic layer nor reconstruct the periodic background inhomogeneities. The justification of this imaging method relies on the well-posedeness of a nonstandard interior transmission problem, which until now was an open problem except for the special case when the local perturbation did not intersect the background inhomogeneities. The analysis of this new interior transmission problem is the main focus of this paper. We then complete the justification of our inversion method and present some numerical examples that confirm the theoretical behavior of the differential indicator function determining the reconstructable regions in the periodic layer.

7.1.3. A robust Expectation-Maximization method for the interpretation of small angle scattering data on dense nanoparticle samples

M. Bakry, H. Haddar and O. Bunau

The Local Monodisperse Approximation (LMA) is a two-parameters model commonly employed for the retrieval of size distributions from the small angle scattering (SAS) patterns obtained on dense nanoparticle samples (e.g. dry powders and concentrated solutions). This work features an original, beyond state-of-the-art implementation of the LMA model resolution for the inverse scattering problem. Our method is based on the Expectation Maximization iterative algorithm and is free from any fine tuning of model parameters. The application of our method on SAS data acquired in laboratory conditions on dense nanoparticle samples is shown to provide very good results.

7.1.4. Detecting Sound Hard Cracks in Isotropic Inhomogeneities

L. Audibert, L. Chesnel, H. Haddar and Kevish Napal

We consider the problem of detecting the presence of sound-hard cracks in a non homogeneous reference medium from the measurement of multi-static far field data. First, we provide a factorization of the far field operator in order to implement the Generalized Linear Sampling Method (GLSM). The justification of the analysis is also based on the study of a special interior transmission problem. This technique allows us to recover the support of the inhomogeneity of the medium but fails to locate cracks. In a second step, we consider a medium with a multiply connected inhomogeneity assuming that we know the far field data at one given frequency both before and after the appearance of cracks. Using the Differential Linear Sampling Method (DLSM), we explain how to identify the component(s) of the inhomogeneity where cracks have emerged. The theoretical justification of the procedure relies on the comparison of the solutions of the corresponding interior transmission problems without and with cracks. Finally we illustrate the GLSM and the DLSM providing numerical results in 2D. In particular, we show that our method is reliable for different scenarios simulating the appearance of cracks between two measurements campaigns

7.1.5. Uncertainty Analysis and Calibration of the Catalytic Properties of Thermal Protection Materials: Formulation of the Bayesian Inference Problem

P.M. Congedo, F. Sanson, T. Magin, F. Panerai

Quantifying the catalytic properties of reusable thermal protection system materials is essential for the design of atmospheric entry vehicles. Their properties quantify the recombination of oxygen and nitrogen atoms into molecules, and allow for accurate computation of the heat flux to the spacecraft. Their rebuilding from ground test data, however, is not straightforward and subject to uncertainties. We propose a fully Bayesian approach to reconstruct the catalytic properties of ceramic matrix composites from sparse high-enthalpy facility experimental data with uncertainty estimates. The results are compared to those obtained by means of an alternative reconstruction procedure, where the experimental measurements are also treated as random variables but propagated through a deterministic solver. For the testing conditions presented in this work, the contribution to the measured heat flux of the molecular recombination is negligible. Therefore, the material catalytic property cannot be estimated precisely. Moreover, epistemic uncertainties are rigorously included, such as the unknown reference calorimeter catalytic property.

7.1.6. A Bayesian framework for the investigation of complex fluid vapor flows

P.M. Congedo, G. Gori, O. Le Maitre, A. Guardone

The present work develops a Bayesian framework for the inference of complex fluid thermodynamic model parameters. The objective is to numerically assess the potential of using experimental measurements to reduce the aleatoric and epistemic uncertainties inherent the Peng-Robinson thermodynamic fluid model for flows of fluids in the non-ideal regimes. Our Bayesian framework is tailored to the design of the TROVA (Test-Rig for Organic VApors) experimental facility, at Politecnico di Milano. Computational Fluid Dynamics (CFD) simulations are used to predict the flow field within the designed test section whereas surrogate models (Polynomial-Chaos expansion) are constructed to account for the predictions dependence on the thermodynamic model parameters. First, synthetic data are generated in the attempt of reproducing a real test case, which is considered as the reference experiment, actually achieved in the TROVA facility. We investigate the resulting posterior uncertainties and assess the knowledge brought by using diverse type of measurements obtained for a flow in the non-ideal regime. Results reveal that the exploitation of pressure measurements only do not allow to infer the thermodynamic coefficients. Indeed, the material-dependent parameters remain highly uncertain.

7.1.7. Shape reconstruction of deposits inside a steam generator using eddy current measurements

H. Girardon, H. Haddar and L. Audibert

Non-destructive testing is an essential tool to assess the safety of the facilities within nuclear plants. In particular, conductive deposits on U-tubes in steam generators constitute a major danger as they may block the cooling loop. To detect these deposits, eddy-current probes are introduced inside the U-tubes to generate currents and measuring back an impedance signal. Based on earlier work on this subject, we develop a shape optimization technique with regularized gradient descent to invert these measurements and recover the deposit shape. To deal with the unknown, and possibly complex, topological nature of the latter, we propose to model it using a level set function. The methodology is first validated on synthetic axisymmetric configurations and fast convergence in ensured by careful adaptation of the gradient steps and regularization parameters. We then consider a more realistic modeling that incorporates the support plate and the presence of imperfections on the tube interior section. We employ in particular an asymptotic model to take into account these imperfections and treat them as additional unknowns in our inverse problem. A multi-objective optimization strategy, based on the use of different operating frequencies, is then developed to solve this problem. Various numerical experimentations with synthetic data demonstrated the viability of our approach.

7.2. Invisiblity and transmission eigenvalues

7.2.1. Trapped modes and reflectionless modes as eigenfunctions of the same spectral problem

A.-S. Bonnet-Ben Dhia, L. Chesnel and V. Pagneux

We consider the reflection-transmission problem in a waveguide with obstacle. At certain frequencies, for some incident waves, intensity is perfectly transmitted and the reflected field decays exponentially at infinity. We show that such reflectionless modes can be characterized as eigenfunctions of an original non-selfadjoint spectral problem. In order to select ingoing waves on one side of the obstacle and outgoing waves on the other side, we use complex scalings (or Perfectly Matched Layers) with imaginary parts of different signs. We prove that the real eigenvalues of the obtained spectrum correspond either to trapped modes (or bound states in the continuum) or to reflectionless modes. Interestingly, complex eigenvalues also contain useful information on weak reflection cases. When the geometry has certain symmetries, the new spectral problem enters the class of $\mathcal{PT}$-symmetric problems.

7.2.2. Transmission eigenvalues with artificial background for explicit material index identification

L. Audibert, L. Chesnel and H. Haddar

We are interested in the problem of retrieving information on the refractive index n of a penetrable inclusion embedded in a reference medium from farfield data associated with incident plane waves. Our approach relies on the use of transmission eigenvalues (TEs) that carry information on n and that can be determined from the knowledge of the farfield operator F. We explain how to modify F into a farfield operator $F^{\mathrm{a}} = F - \widetilde{F}$, where $\widetilde{F}$ is computed numerically, corresponding to well chosen artificial background and for which the associated TEs provide more accessible information on n.

7.2.3. Simple examples of perfectly invisible and trapped modes in waveguides

L. Chesnel and V. Pagneux

We consider the propagation of waves in a waveguide with Neumann boundary conditions. We work at low wavenumber focusing our attention on the monomode regime. We assume that the waveguide is symmetric with respect to an axis orthogonal to the longitudinal direction and is endowed with a branch of height L whose width coincides with the wavelength of the propagating modes. In this setting, tuning the parameter L, we prove the existence of simple geometries where the transmission coefficient is equal to one (perfect invisibility). We also show that these geometries, for possibly different values of L, support so called trapped modes (non zero solutions of finite energy of the homogeneous problem) associated with eigenvalues embedded in the continuous spectrum.

7.2.4. New sets of eigenvalues in inverse scattering for inhomogeneous media and their determination from scattering data

F. Cakoni, H. Haddar and L. Audibert

We developed a general mathematical framework to determine interior eigenvalues from a knowledge of the modified far field operator associated with an unknown (anisotropic) inhomogeneity. The modified far field operator is obtained by subtracting from the measured far field operator the computed far field operator corresponding to a well-posed scattering problem depending on one (possibly complex) parameter. Injectivity of this modified far field operator is related to an appropriate eigenvalue problem whose eigenvalues can be determined from the scattering data, and thus can be used to obtain information about material properties of the unknown inhomogeneity. We discuss here two examples of such modification leading to a Steklov eigenvalue problem, and a new type of the transmission eigenvalue problem. We present some numerical examples demonstrating the viability of our method for determining the interior eigenvalues form far field data.

7.2.5. The Asymptotic of Transmission Eigenvalues for a Domain with a Thin Coating

H. Boujlida, H Haddar and M. Khenissi

We consider the transmission eigenvalue problem for a medium surrounded by a thin layer of inhomogeneous material with different refractive index. We derive explicit asymptotic expansion for the transmission eigenvalues with respect to the thickness of the thin layer. We prove error estimate for the asymptotic expansion up to order 1 for simple eigenvalues. This expansion can be used to obtain explicit expressions for constant index of refraction.

7.2.6. The spectral analysis of the interior transmission eigenvalue problem for Maxwell's equations

H. Haddar and S. Meng

we consider the transmission eigenvalue problem for Maxwell's equations corresponding to non-magnetic inhomogeneities with contrast in electric permittivity that has fixed sign (only) in a neighborhood of the boundary. Following the analysis made by Robbiano in the scalar case we study this problem in the framework of semiclassical analysis and relate the transmission eigenvalues to the spectrum of a Hilbert-Schmidt operator. Under the additional assumption that the contrast is constant in a neighborhood of the boundary, we prove that the set of transmission eigenvalues is discrete, infinite and without finite accumulation points. A notion of generalized eigenfunctions is introduced and a denseness result is obtained in an appropriate solution space.

7.2.7. Non reflection and perfect reflection via Fano resonance in waveguides

L. Chesnel, S.A. Nazarov

We investigate a time-harmonic wave problem in a waveguide. By means of asymptotic analysis techniques, we justify the so-called Fano resonance phenomenon. More precisely, we show that the scattering matrix considered as a function of a geometrical parameter ε and of the frequency λ is in general not continuous at a point $(\varepsilon, \lambda) = (0, \lambda^0)$ where trapped modes exist. In particular, we prove that for a given $\varepsilon \neq 0$ small, the scattering matrix exhibits a rapid change for frequencies varying in a neighbourhood of λ^0. We use this property to construct examples of waveguides such that the energy of an incident wave propagating through the structure is perfectly transmitted (non reflection) or perfectly reflected in monomode regime. We provide numerical results to illustrate our theorems.

7.2.8. From zero transmission to trapped modes in waveguides

L. Chesnel, V. Pagneux

We consider a time-harmonic scattering wave problem in a 2D waveguide at wavenumber k such that one mode is propagating in the far field. For a given k, playing with one scattering branch of finite length, we demonstrate how to construct geometries with zero transmission. The main novelty in this result is that the symmetry of the geometry is not needed: the proof relies on the unitary structure of the scattering matrix. Then, from a waveguide with zero transmission, we show how to build geometries supporting trapped modes associated with eigenvalues embedded in the continuous spectrum. For this second construction, using the augmented scattering matrix and its unitarity, we play both with the geometry and the wavenumber. The mathematical analysis is supplemented by numerical illustrations of the results.

7.3. Shape and topology optimization

7.3.1. Taking into account thermal residual stresses in topology optimization of structures built by additive manufacturing

G. Allaire and L. Jakabcin.

We introduce a model and several constraints for shape and topology optimization of structures, built by additive manufacturing techniques. The goal of these constraints is to take into account the thermal residual stresses or the thermal deformations, generated by processes like Selective Laser Melting, right from the beginning of the structural design optimization. In other words, the structure is optimized concurrently for its final use and for its behavior during the layer by layer production process. It is well known that metallic additive manufacturing generates very high temperatures and heat fluxes, which in turn yield thermal deformations that may prevent the coating of a new powder layer, or thermal residual stresses that may hinder the mechanical properties of the final design. Our proposed constraints are targeted to avoid these undesired effects. Shape derivatives are computed by an adjoint method and are incorporated into a level set numerical optimization algorithm. Several 2-d and 3-d numerical examples demonstrate the interest and effectiveness of our approach.

7.3.2. Topology optimization of modulated and oriented periodic microstructures by the homogenization method

G. Allaire, P. Geoffroy-Donders and O. Pantz

This work is concerned with the topology optimization of structures made of periodically perforated material, where the microscopic periodic cell can be macroscopically modulated and oriented. The main idea is to optimize the homogenized formulation of this problem, which is an easy task of parametric optimization, then to project the optimal microstructure at a desired lengthscale, which is a delicate issue, albeit computationally cheap. The main novelty of our work is, in a plane setting, the conformal treatment of the optimal orientation of the microstructure. In other words, although the periodicity cell has varying parameters and orientation throughout the computational domain, the angles between its members or bars are conserved. The main application of our work is the optimization of so-called lattice materials which are becoming increasingly popular in the context of additive manufacturing. Several numerical examples are presented for single and multiple loads problems, as well as for compliance or more general objective functions.

7.3.3. Shape optimization of a coupled thermal fluid-structure problem in a level set mesh evolution framework

G. Allaire, F. Feppon, F. Bordeu, J. Cortial and C. Dapogny

Hadamard's method of shape differentiation is applied to topology optimization of a weakly coupled three physics problem. The coupling is weak because the equations involved are solved consecutively, namely the steady state Navier-Stokes equations for the fluid domain, first, the convection diffusion equation for the whole domain, second, and the linear thermo-elasticity system in the solid domain, third. Shape sensitivities are derived in a fully Lagrangian setting which allows us to obtain shape derivatives of general objective functions. An emphasis is given on the derivation of the adjoint interface condition dual to the one of equality of the normal stresses at the fluid solid interface. The arguments allowing to obtain this surprising condition are specifically detailed on a simplified scalar problem. Numerical test cases are presented using a level set mesh evolution method. It is demonstrated how the implementation enables to treat a variety of shape optimization problems.

7.3.4. Optimizing supports for additive manufacturing

G. Allaire and B. Bogosel

In additive manufacturing process support structures are often required to ensure the quality of the final built part. In this article we present mathematical models and their numerical implementations in an optimization loop, which allow us to design optimal support structures. Our models are derived with the requirement that they should be as simple as possible, computationally cheap and yet based on a realistic physical modeling. Supports are optimized with respect to two different physical properties. First, they must support overhanging regions of the structure for improving the stiffness of the supported structure during the building process. Second, supports can help in channeling the heat flux produced by the source term (typically a laser beam) and thus improving the cooling down of the structure during the fabrication process. Of course, more involved constraints or manufacturability conditions could be taken into account, most notably removal of supports. Our work is just a first step, proposing a general framework for support optimization. Our optimization algorithm is based on the level set method and on the computation of shape derivatives by the Hadamard method. In a first approach, only the shape and topology of the supports are optimized, for a given and fixed structure. In second and more elaborated strategy, both the supports and the structure are optimized, which amounts to a specific multiphase optimization problem. Numerical examples are given in 2-d and 3-d.

7.3.5. Structural optimization under internal porosity constraints using topological derivatives

G. Allaire, J.Martinez-Frutos, C. Dapogny, F. Periago

Porosity is a well-known phenomenon occurring during various manufacturing processes (casting, welding, additive manufacturing) of solid structures, which undermines their reliability and mechanical performance. The main purpose of this article is to introduce a new constraint functional of the domain which controls the negative impact of porosity on elastic structures in the framework of shape and topology optimization. The main ingredient of our modelling is the notion of topological derivative, which is used in a slightly unusual way: instead of being an indicator of where to nucleate holes in the course of the optimization process, it is a component of a new constraint functional which assesses the influence of pores on the mechanical performance of structures. The shape derivative of this constraint is calculated and incorporated into a level set based shape optimization algorithm. Our approach is illustrated by several two- and three-dimensional numerical experiments of topology optimization problems constrained by a control on the porosity effect.

7.4. Analysis of some wave problems

7.4.1. On well-posedness of time-harmonic problems in an unbounded strip for a thin plate model

L. Bourgeois, L. Chesnel, S. Fliss

We study the propagation of elastic waves in the time-harmonic regime in a waveguide which is unbounded in one direction and bounded in the two other (transverse) directions. We assume that the waveguide is thin in one of these transverse directions, which leads us to consider a Kirchhoff-Love plate model in a locally perturbed 2D strip. For time harmonic scattering problems in unbounded domains, well-posedness does not hold in a classical setting and it is necessary to pre- scribe the behaviour of the solution at infinity. This is challenging for the model that we consider and constitutes our main contribution. Two types of boundary conditions are considered: either the strip is simply supported or the strip is clamped. The two boundary conditions are treated with two different methods. For the simply supported problem, the analysis is based on a result of Hilbert basis in the transverse section. For the clamped problem, this property does not hold. Instead we adopt the Kondratiev's approach, based on the use of the Fourier transform in the unbounded direction, together with techniques of weighted Sobolev spaces with detached asymptotics. After introducing radiation conditions, the corresponding scattering problems are shown to be well-posed in the Fredholm sense. We also show that the solutions are the physical (outgoing) solutions in the sense of the limiting absorption principle.

7.4.2. Crime Pays: Homogenized Wave Equations for Long Times

G. Allaire, A. Lamacz and J. Rauch

This work examines the accuracy for large times of asymptotic expansions from periodic homogenization of wave equations. As usual, ϵ denotes the small period of the coefficients in the wave equation. We first prove that the standard two scale asymptotic expansion provides an accurate approximation of the exact solution for times t of order $\epsilon^{-2+\delta}$ for any $\delta > 0$. Second, for longer times, we show that a different algorithm, that is called criminal because it mixes different powers of ϵ, yields an approximation of the exact solution with error $O(\epsilon^N)$ for times ϵ^{-N} with N as large as one likes. The criminal algorithm involves high order homogenized equations that, in the context of the wave equation, were first proposed by Santosa and Symes and analyzed by Lamacz. The high order homogenized equations yield dispersive corrections for moderate wave numbers. We give a systematic analysis for all time scales and all high order corrective terms.

7.5. Diffusion MRI

7.5.1. A partition of unity finite element method for computational diffusion MRI

D. V. Nguyen, J. Jansson, J. Hoffman and J.-R. Li.

The Bloch-Torrey equation describes the evolution of the spin (usually water proton) magnetization under the influence of applied magnetic field gradients and is commonly used in numerical simulations for diffusion MRI and NMR. Microscopic heterogeneity inside the imaging voxel is modeled by interfaces inside the simulation domain, where a discontinuity in the magnetization across the interfaces is produced via a permeability coefficient on the interfaces. To avoid having to simulate on a computational domain that is the size of an entire imaging voxel, which is often much larger than the scale of the microscopic heterogeneity as well as the mean spin diffusion displacement, smaller representative volumes of the imaging medium can be used as the simulation domain. In this case, the exterior boundaries of a representative volume either must be far away from the initial positions of the spins or suitable boundary conditions must be found to allow the movement of spins across these exterior boundaries.

Many approaches have been taken to solve the Bloch-Torrey equation but an efficient high performance computing framework is still missing. In this paper, we present formulations of the interface as well as the exterior boundary conditions that are computationally efficient and suitable for arbitrary order finite elements and parallelization. In particular, the formulations are based on the partition of unity concept which allows for a discontinuous solution across interfaces conforming with the mesh with weak enforcement of real (in the case of interior interfaces) and artificial (in the case of exterior boundaries) permeability conditions as well as an operator splitting for the exterior boundary conditions. The method is straightforward to implement and it is available in FEniCS for moderate-scale simulations and in FEniCS-HPC for large-scale simulations. The order of accuracy of the resulting method is validated in numerical tests and a good scalability is shown for the parallel implementation. We show that the simulated dMRI signals offer good approximations to reference signals in cases where the latter are available and we performed simulations for a realistic model of a neuron to show that the method can be used for complex geometries.

7.5.2. Diffusion MRI simulation in thin-layer and thin-tube media using a discretization on manifolds

D. V. Nguyen, J. Jansson, H. T. A. Tran, J. Hoffman and J.-R. Li.

The Bloch-Torrey partial differential equation describes the evolution of the transverse magnetization of the imaged sample under the influence of diffusion-encoding magnetic field gradients inside the MRI scanner. The integral of the magnetization inside a voxel gives the simulated diffusion MRI signal. This paper proposes a finite element discretization on manifolds in order to simulate the diffusion MRI signal in domains that have a thin layer or a thin tube geometrical structure. Suppose that the three-dimensional domain has a thin layer structure: points in the domain can be obtained by starting on the two-dimensional manifold and moving along a depth (thickness) function. For this type of domains, we propose a finite element discretization formulated on a surface triangulation of the manifold. The variable thickness of the domain is included in the weak formulation on the surface triangular elements. A simple modification extends the approach to ‘thin tube’ domains where a manifold in one dimension and a two-dimensional variable cross-section describe the points in the domain. We conducted a numerical study of the proposed approach by simulating the diffusion MRI signals from the extracellular space (a thin layer medium) and from neurons (a thin tube medium), comparing the results with the reference signals obtained using a standard three-dimensional finite element discretization. We show good agreement between the simulated signals using our proposed method and the reference signals. The approximation becomes better as the diffusion time increases. The method helps to significantly reduce the required simulation time, computational memory, and difficulties associated with mesh generation, thus opening the possibilities to simulating complicated structures at low cost for a better understanding of diffusion MRI in the brain.

7.5.3. The time-dependent diffusivity in the abdominal ganglion of Aplysia californica, experiments and simulations

K. V. Nguyen, D. Le Bihana, L. Ciobanua and J.-R. Li

The nerve cells of the *Aplysia* are much larger than mammalian neurons. Using the *Aplysia* ganglia to study the relationship between the cellular structure and the diffusion MRI signal can potentially shed light on this relationship for more complex organisms. We measured the dMRI signal of chemically-fixed abdominal ganglia of the *Aplysia* at several diffusion times. At the diffusion times measured and observed at low b-values, the dMRI signal is mono-exponential and can be accurately represented by the parameter ADC (Apparent Diffusion Coefficient).

We performed numerical simulations of water diffusion for three types of cells in the abdominal ganglia: the large cell neurons, the bag cells, and the nerve cells. For the bag cells and nerves cells, we created spherical and cylindrical geometrical configurations that are consistent with known information about the cellular structures from the literature. We used the simulation results to obtain information about the intrinsic diffusion coefficient in these cells.

For the large cell neurons, we created geometrical configurations by segmenting high resolution T_2-weighted (T_2w) images to obtain the cell outline and then incorporated a manually generated nucleus. We used numerical simulations to validate the claim that water diffusion in the large cell neurons is in the short diffusion time regime for our experimental diffusion times.

Then, using the analytical short time approximation (STA) formula for the ADC, we showed that in order to explain the experimentally observed behavior in the large cell neurons, it is necessary to consider the nucleus and the cytoplasm as two separate diffusion compartments. By using a two compartment STA model, we were able to illustrate the effect of the highly irregular shape of the cell nucleus on the ADC.

7.5.4. The derivation of homogenized diffusion kurtosis models for diffusion MRI

H. Haddar, M. Kchaou and M. Moakher

We use homogenization theory to establish a new macroscopic model for the complex transverse water proton magnetization in a voxel due to diffusion-encoding magnetic field gradient pulses in the case of biological tissue with impermeable membranes. In this model, new higher-order diffusion tensors emerge and offer more information about the structure of the biological tissues. We explicitly solve the macroscopic model to obtain an ordinary differential equation for the diffusion MRI signal that has similar structure as diffusional kurtosis imaging models. We finally present some validating numerical results on synthetic examples showing the accuracy of the model with respect to signals obtained by solving the Bloch-Torrey equation.

7.5.5. On-going collaborative projects on DMRI

J.R. Li, H. Haddar and I. Mekkaoui

- We performed simulations for a collaborative project with Demian Wassermann of the Parietal team on distinguishing between Spindle and pyramidal neurons with Multi-shell Diffusion MRI.
- We continue in the simulation and modeling of heart diffusion MRI with the post-doc project of Imen Mekkaoui, funded by Inria-EPFL lab. The project is co-supervised with Jan Hesthaven, Chair of Computational Mathematics and Simulation Science (MCSS), EPFL.

7.6. Mathematical tools for Psychology

J. R. Li and J. Hao

This is the start of a collaborative effort between the Defi team and Dr. Hassan Rahioui at the centre hospitalier Sainte Anne and l'Université Paris Diderot.

- We started a new research direction in algorithm and software development for analysis and classification of EEG measurements during the administration of neuropsychological tests for AD/HD with the PhD project of Jingjing Hao, co-supervised with Dr. Hassan Rahioui, Chef du pôle psychiatrique du 7e arrondissement de Paris rattaché au centre hospitalier Sainte-Anne.

  Result: unfortunately Jingjing Hao will not be able to continue with this PhD project as of Jan 2019. We will modify the project in consultation with Dr. Rahioui and continue it in another format.

7.7. Shape optimization under uncertainties

7.7.1. Surrogate-Assisted Bounding-Box Approach Applied to Constrained Multi-Objective Optimisation Under Uncertainty

P.M. Congedo, M. Rivier

This work is devoted to tackling constrained multi-objective optimisation under uncertainty problems. In particular, the SABBa (Surrogate-Assisted Bounding-Box approach) framework is applied and extended to handle both robust and reliability-based constrained optimisation problems. This approach aims at efficiently dealing with uncertainty-based optimisation problems, with approximated robustness and reliability measures. A Bounding-Box (or conservative box) is defined as a multi-dimensional product of intervals centred on approximated objectives and constraints and containing the underlying true values. In SABBa, this approach is supplemented with a Surrogate-Assisting strategy, which is very effective to reduce the overall computational cost, notably during the last iterations of the optimisation. The efficiency of the method is further increased using the concept of Pareto Optimal Probability (POP) computed for each box, and proposing some estimations for conservative error computation and box refinement using a Gaussian Process (GP).

7.7.2. *A quantile-based optimization under uncertainty of an ORC turbine cascade*

P.M. Congedo, N. Razaaly

This study presents an original and fast robust shape optimization approach to overcome the limitation of a deterministic optimization that neglects operating conditions variability, applied on a typical 2D ORC turbine cascade (Biere). Flow around the blade is solved by means of inviscid simulation using the open-source SU2 code, considering Non-Ideal gas effects modeled through the use of the Peng-Robinson-Stryjek-Vera equation of state, from which a Quantity of Interest (QoI) is recovered. We propose here a mono-objective formulation consisting in minimizing the α-quantile of the QoI under a constraint, at a low computational cost. This is performed by using an efficient robust optimization approach, coupling a state-of- the-art quantile estimation and a classical bayesian optimization method. First, the advantages of a quantile-based formulations are illustrated with respect to a classical mean-based robust optimization. Secondly, we demonstrate the effectiveness of applying this robust optimization framework with a low-fidelity inviscid solver by comparing the resulting optimal design with the ones obtained with a deterministic optimization using a high-fidelity turbulent solver.

7.8. Uncertainty Quantification methods for uncertainty propagation

7.8.1. *Kriging-sparse Polynomial Dimensional Decomposition surrogate model with adaptive refinement*

P.M. Congedo, A. Cortesi, G. El Jannoun

In this work, an algorithm for the construction of a low-cost and accurate metamodel is proposed, having in mind redcomputationally expensive applications. It has two main features. First, Universal Kriging is coupled with sparse Polynomial Dimensional Decomposition (PDD) to build a metamodel with improved accuracy. The polynomials selected by the adaptive PDD representation are used as a sparse basis to build a Universal Kriging surrogate model. Secondly, a numerical method, derived from anisotropic mesh adaptation, is formulated in order to adaptively insert a fixed number of new training points to an existing Design of Experiments. The convergence of the proposed algorithm is analyzed and assessed on different test functions with an increasing size of the input space. Finally, the algorithm is used to propagate uncertainties in two high-dimensional real problems related to the atmospheric reentry.

7.8.2. *Novel algorithm using Active Metamodel Learning and Importance Sampling: Application to multiple failure regions of low probability*

P.M. Congedo, N. Razaaly

Calculation of tail probabilities is of fundamental importance in several domains, such as in risk assessment. One major challenge consists in the computation of low-failure probability in cases characterized by multiple-failure regions, especially when an unbiased estimation of the error is required. Methods developed in literature rely mostly on the construction of an adaptive surrogate, tackling some problems such as the metamodel building criterion and the global computational cost, at the price of a generally biased estimation of the failure probability. In this work, we propose a novel algorithm suitable for low-failure probability and multiple-failure regions, permitting to both building an accurate metamodel and to provide a statistically consistent error. Indeed, an importance sampling technique is used, which is quasi-optimal since permits, by exploiting the knowledge of the metamodel, to provide two unbiased estimators of the failure probability. Additionally, a gaussian mixture-based importance sampling technique is proposed, permitting to drastically reduce the computational cost when estimating some reference values, or the failure probability directly from the metamodel. Several numerical examples are carried out, showing the very good performances of the proposed method with respect to the state-of-the-art in terms of accuracy and computational cost. A physical test-case, focused on the numerical simulation of non-ideal gas turbine cascades, is also investigated to illustrate the capabilities of the method on an industrial case.

7.8.3. Uncertainty propagation framework for systems of codes

P.M. Congedo, F. Sanson, O. Le Maitre

The simulation of complex multi-physics phenomena often requires the use of coupled solvers, modelling different physics (fluids, structures, chemistry, etc) with largely differing computational complexities. We call Systems of Solvers (SoS) a set of interdependent solvers where the output of an upstream solver can be the input of a downstream solvers. In this work we restrict ourselves to weakly coupled problems. A system of solvers typically encapsulate a large number of uncertain input parameters, challenging classical Uncertainty Quantification (UQ) methods such as spectral expansions and Gaussian process models which are affected by the curse of dimensionality. In this work, we develop an original mathematical framework, based on Gaussian Processes (GP) to construct a global metamodel of the uncertain SoS that can be used to solve forward and backward UQ problems. The key idea of the proposed approach is to determine a local GP model for each solver of the SoS. These local GP models are built adaptively to satisfy criteria based on the global output error estimation, which can be decomposed (following an ANOVA-like decomposition) into contributions from individual GP models. This decomposition enables one to select the local GP models that need be refined to efficiently reduce the global error using computer experiment design methods or Bayesian optimization. This framework is then applied to a space object reentry problem.

7.9. Application of Uncertainty Quantification studies to fluid-dynamics problems

7.9.1. Validation of the Non-Ideal Compressible-Fluid Dynamics solver from the open-source SU2 suite

P.M. Congedo, G. Gori, A. Guardone, M. Zocca

The first-ever experimental validation of a flow simulation software for Non-Ideal Compressible-Fluid Dynamics (NICFD) flows is presented. Numerical results from the open-source suite SU2 are compared against pressure and Mach number measurements of supersonic flows of siloxane fluid MDM (Octamethyltrisiloxane, $C_8H_{24}O_2Si_3$) at conditions in the close proximity of the liquid-vapour saturation curve. The test set is representative of typical operating conditions of Organic Rankine Cycle systems and it includes expanding flows through a converging-diverging nozzle in mildly-to-highly non-ideal conditions. The validation process takes advantage of an Uncertainty Quantification analysis, to estimate the variability of the numerical solution with respect to the physical uncertainties and to provide a robust assessment of the SU2 capabilities. All considered flows are well represented by the numerical solutions and therefore the reliability of the numerical implementation and the predictiveness of the NICFD solver are confirmed.

7.9.2. Impact of geometric, operational, and model uncertainties on the non-ideal flow through a supersonic ORC turbine cascade

P.M. Congedo, N. Razaaly, G. Persico

Typical energy sources for Organic Rankine Cycle (ORC) power systems feature variable heat load and turbine inlet/outlet thermodynamic conditions. The use of organic compounds with heavy molecular weight introduces uncertainties in the fluid thermodynamic modeling. In addition, the peculiarities of organic fluids typically lead to supersonic turbine configurations featuring supersonic flows and shocks, which grow in relevance in the aforementioned off-design conditions; these features also depend strongly on the local blade shape, which can be influenced by the geometric tolerances of the blade manufacturing. This study presents an Uncertainty Quantification (UQ) analysis on a typical supersonic nozzle cascade for ORC applications, by considering a two-dimensional high-fidelity turbulent Computational Fluid Dynamic (CFD) model. Kriging-based techniques are used in order to take into account at a low computational cost, the combined effect of uncertainties associated to operating conditions, fluid parameters, and geometric tolerances. The geometric variability is described by a finite Karhunen-Loeve expansion representing a non-stationary Gaussian random field, entirely defined by a null mean and its autocorrelation function. Several results are illustrated about the ANOVA decomposition of several quantities of interest for different operating conditions, showing the importance of geometric uncertainties on the turbine performances.

7.9.3. Efficient surrogate based human risk estimation of a space object reentry

P.M. Congedo, F. Sanson, O. Le Maitre, J.-M. Bouilly, C. Bertorello

The prediction of risk associated with the reentry of a man made space object is critical but subject to input parameter uncertainties. To compute the risk one needs to determine whether the object survives to reentry and if it does where it falls on Earth. Expensive numerical models can be used to answer both questions but they can only be evaluated a limited number of times to propagate the uncertainties. In this work, we present an original approach to construct an accurate surrogate model of the numerical models using a limited number of solver evaluations. Using Gaussian Processes, the constructed surrogate model is able two answer both questions (survivability and impact location) in order to provide an accurate description of the risk. The surrogate model can achieve high level of accuracy in terms of risk estimation using dedicated active learning strategies. The efficiency of the method is illustrated on analytical test cases and an actual space object reentry case.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- A CIFRE PhD thesis started in December 2015 with Safran Tech. The student is Mrs Perle Geoffroy who is working on "topology optimization by the homogenization method in the context of additive manufacturing".
- A CIFRE PhD thesis started in April 2017 with Safran Tech. The student is M. Florian Feppon who is working on "topology optimization for a coupled thermal-fluid-structure system”.
- A CIFRE PhD thesis started in October 2017 with Renault. The student is Mrs Lalaina Rakotondrainibe who is working on "topology optimization of connections between mechanical parts”.
- A CIFRE PhD thesis started November 2017 with EDF. The student is H. Girardon who is working on "level set method for eddy current non destructive testting".
- A CIFRE PhD thesis started May 2017 with ArianeGroup. The student is M. Mickael Rivier who is working on "Optimization under uncertainty methods for expensive computer codes".
- A CIFRE PhD thesis started November 2018 with CEA CESTA. The student is M. Paul Novello who is working on "Deep Learning for atmospheric reentry".

8.2. Bilateral Grants with Industry

- The SOFIA project (SOlutions pour la Fabrication Industrielle Additive métallique) started in the summer of 2016. Its purpose is to make research in the field of metallic additive manufacturing. The industrial partners include Michelin, FMAS, ESI, Safran and others. The academic partners are different laboratories of CNRS, including CMAP at Ecole Polytechnique. The project is funded for 6 years by BPI (Banque Publique d'Investissement).
- G. Allaire is participating to the TOP project at IRT SystemX which started in February 2017. It is concerned with the development of a topology optimization platform with industrial partners (Renault, Safran, Airbus, ESI).
- FUI project Saxsize. This three years project started in October 2015 and extended till April 2019 and it involves Xenocs (coordinator), Inria (DEFI), Pyxalis, LNE, Cordouan and CEA. It is a followup of Nanolytix where a focus is put on SAXS quantifications of dense nanoparticle solutions.
- Contract with THALES, Activity around the numerical certification of debris codes, Coordinator: P.M. Congedo.
- Contract with ArianeGroup, Activity around techniques for Uncertainty Quantification, Coordinator: P.M. Congedo.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. EVE

- Title : Virtual prototyping of EVE engines
- Type : Co-funded from Region Aquitaine and Inria
- Duration : 36 months
- Starting : October 2018
- Coordinator : P.M. Congedo
- Abstract : The main objective of this thesis is the construction of a numerical platform, for permitting an efficient virtual prototyping of the EVE expander. This will provide EXOES with a numerical tool, that is much more predictive with respect to the tools currently available and used in EXOES, by respecting an optimal trade-off in terms of complexity/cost needed during an industrial design process.i Two research axes will be mainly developed. First, the objective is to perform some high-predictive numerical simulation for reducing the amount of experiments, thanks to a specific development of RANS tools (Reynolds Averaged Navier-Stokes equations) for the fluids of interest for EXOES. These tools would rely on complex thermodynamic models and a turbulence model that should be modified. The second axis is focused on the integration of the solvers of different fidelity in a multi-fidelity platform for performing optimization under uncertainties. The idea is to evaluate the system performances by using massively the low-fidelity models, and by correcting these estimations via only few calculations with the high-fidelity code.

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

9.2.1.1. UTOPIAE

Program: H2020 MSCA-ITN

Project acronym: UTOPIAE

Project title: Handling the unknown at the edge of tomorrow

Duration: January 2017- December 2020

Coordinator: M. Vasile (Strathclyde University)

Other partners: see http://utopiae.eu/ for additional details

UTOPIAE is a European research and training network looking at cutting edge methods bridging optimisation and uncertainty quantification applied to aerospace systems. The network will run from 2017 to 2021, and is funded by the European Commission through the Marie Skłodowska-Curie Actions of H2020. The network is made up of 15 partners across 6 European countries, including the UK, and one international partner in the USA, collecting mathematicians, engineers and computer scientists from academia, industry, public and private sectors.

Mission statement : To train, by research and by example, 15 Early Stage Researchers in the field of uncertainty quantification and optimisation to become leading independent researchers and entrepreneurs that will increase the innovation capacity of the EU. To equip the researchers with the skills they will need for successful careers in academia and industry. To develop fundamental mathematical methods and algorithms to bridge the gap between Uncertainty Quantification and Optimisation and between Probability Theory and Imprecise Probability Theory for Uncertainty Quantification to efficiently solve high-dimensional, expensive and complex engineering problems.

9.3. International Initiatives

9.3.1. Inria International Labs

P.M. Congedo is the Inria Coordinator of the CWI-Inria Inria International Lab.

IIL CWI-Inria

Associate Team involved in the International Lab:

9.3.1.1. COMMUNES

Title: Computational Methods for Uncertainties in Fluids and Energy Systems

International Partner (Institution - Laboratory - Researcher):

CWI (Netherlands) - Scientific Computing Group - Daan Crommelin

Start year: 2017

See also: https://project.inria.fr/inriacwi/projects/communes/

This project aims to develop numerical methods capable to take into account efficiently unsteady experimental data, synthetic data coming from numerical simulation and the global amount of uncertainty associated to measurements, and physical-model parameters. We aim to propose novel algorithms combining data-inferred stochastic modeling, uncertainty propagation through computer codes and data assimilation techniques. The applications of interest are both related to the exploitation of renewable energy sources: wind farms and solar Organic Rankine Cycles (ORCs).

9.3.1.2. Informal International Partners

University of Zurich : R. Abgrall. Collaboration on high order adaptive methods for CFD and uncertainty quantification.

Politecnico di Milano, Aerospace Department (Italy) : Pr. A. Guardone. Collaboration on ALE for complex flows (compressible flows with complex equations of state).

von Karman Institute for Fluid Dynamics (Belgium). With Pr. T. Magin we work on Uncertainty Quantification problems for the identification of inflow condition of hypersonic nozzle flows.

Rutgers University. Collaboration with Pr. F. Cakoni on transmission eigenvalues.

University of Delaware. Collaboration with Pr. D. Colton on inverse scattering theory

Ecole Nationale des Ingénieurs de Tunis. Collaboration with Pr. M. Moakher on Diffusion MRI

9.4. International Research Visitors

9.4.1. Visits of International Scientists

- Fioralba Cakoni, one month, July 15-August 14, 2018

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. General Chair, Scientific Chair

- P. Congedo is General Chair of the CWI-Inria workshop at Inria Paris research centre in Paris on September 25, 26 2018.

10.1.2. Member of the Organizing Committees

- L. Chesnel co-organized the Journée de rentrée (2018) of the Centre de Mathématiques Appliquées of École Polytechnique.
- L. Chesnel co-organizes the seminar of the Centre de Mathématiques Appliquées of École Polytechnique.
- L. Chesnel co-organizes the joint seminar of the Inria teams Defi-M3DISIM-Poems.
- H. Haddar co-organized a minisymposium at waves conference, Karlsruhe, July 2018
- H. Haddar co-organized a minisymposium at the conference Inverse problems: modeling and simulation, Malta, May 2018
- H. haddar co-organized a minisymposium at the Faculty of DSIT of Ecole polytechnique, May 2018.
- J.R. Li is Co-organizer of the summer school Ecole d'Été France Excellence, Data science for document analysis and understanding sponsored by the French Embassy in China, 07/2018. 4 weeks.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- G. Allaire is member of the editorial boards of
  - Book series "Mathématiques et Applications" of SMAI and Springer,
  - ESAIM/COCV, Structural and Multidisciplinary Optimization,
  - Discrete and Continuous Dynamical Systems Series B,
  - Computational and Applied Mathematics,
  - Mathematical Models and Methods in Applied Sciences (M3AS),
  - Annali dell'Universita di Ferrara,
  - OGST (Oil and Gas Science and Technology),
  - Journal de l'Ecole Polytechnique - Mathématiques,
  - Journal of Optimization Theory and Applications.
- P.M. Congedo is Editor of Mathematics and Computers in Simulation, MATCOM (Elsevier).
- H. Haddar is
  - Member the editorial board of Inverse Problems
  - Associate Editor of the SIAM Journal on Scientific Computing
  - Guest editor of for a special issue in the journal Inverse Problems

10.1.3.2. Reviewer - Reviewing Activities

We reviewed papers for top international journals in the main scientific themes of the team.

10.1.4. Invited Talks

- G. Allaire
  - Seminar at BCAM, Bilbao (january 2018).
  - Séminaire JOFA, Pau (juin 2018).
  - ECCM-ECFD, ECCOMAS, Glasgow (juin 2018).
  - Summer school Sendai, Japon (August 2018).
  - Fifth workshop on thin structures, Naples, September 13-15, 2018.
  - Current trends and open problems in computational solid mechanics, Hannover, October 8-9, 2018.
- L. Chesnel
  - Conference on Mathematics of Wave Phenomena, July 2018.
  - SIAM conference on imaging science, Bologna, June 2018.
  - Inverse problems: modeling and simulation, Malta, May 2018.
  - Workshop du GDR ondes, Jussieu, March 2018.
- P.M. Congedo
  - von Karman Institute Symposium, Bruxelles, April 2018.
  - CERFACS Seminar, Toulouse, October 2018
  - VKI Lecture Series on Uncertainty Quantification, Bruxelles, September 2018.
  - Seminar at CEA-CESTA, Le Barp, September 2018.
- H. Haddar
  - ICAV conference, Hammamet, March 2018
  - SIAM conference on imaging science, Bologna, June 2018.
  - Waves conference, Karsruhe, July 2018
  - Seminar at LJK, Grenoble, October 2018
  - Kickoff workshop of Mecawave, November 2018
  - Workshop "Inverse Problems: Theory and Applications", Reims, 2018

10.1.5. Leadership within the Scientific Community

- G. Allaire is a board member of Institut Henri Poincaré (IHP). He is the chairman of the scientific council of IFPEN (French Petroleum Institute and New Energies). He is the chairman of the scientific council of AMIES (Agency for Interaction in Mathematics with Business and Society).

10.1.6. Scientific Expertise

- G. Allaire is a member of th"comité national" CNRS, section 41 (mathematics).
- G. Allaire is a member of the scientific board of the Gaspard Monge program on optimization (PGMO) at the Jacques Hadamard Mathematical Foundation.
- J.R. Li is Member of the SIAM Committee on Programs and Conferences 2017-2019.
- J.R. Li is Member Elu of Inria Commission d'Evaluation, 2015-present.

10.1.7. Research Administration

- J.R. Li is correspondant International for Centre de Mathematiques Appliquees, Ecole Polytechnique, 2018-present.
- J.R. Li is responsable for the Ecole Polytechnique part of the French-Vietnam Master Program in Applied Mathematics, 2016-present.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- Master: Grégoire Allaire, Approximation Numérique et Optimisation, for students in the second year of Ecole Polytechnique curriculum: 8 lessons of 1h30.
- Master: Grégoire Allaire, Transport and diffusion, for students in the third year of Ecole Polytechnique curriculum. 9 lessons of 2h jointly with F. Golse.
- Master: Houssem Haddar, Approximation Numérique et Optimisation, for students in the second year of Ecole Polytechnique curriculum: 8 TDs of 4h.
- Master: Houssem Haddar, Waves and imaging: Concepts, Theory and Applications, Master M2 ”mathematical modeling”: 9 lessons of 3h.
- Master: Lucas Chesnel, Variational analysis for partial differential equations, for students in the second year of Ecole Polytechnique curriculum: 8 TDs of 4h.
- Master: Lucas Chesnel, Numerical approximation and optimisation, for students in the second year of Ecole Polytechnique curriculum: 2 TDs of 4h + one project.
- Master: Lucas Chesnel, Modal Modélisation mathématique par la démarche expérimentale, for students in the second year of Ecole Polytechnique curriculum: 5 TDs of 2h.
- Master: Grégoire Allaire, Optimal design of structures, for students in the third year of Ecole Polytechnique curriculum. 9 lessons of 1h30.
- Master: Grégoire Allaire, Theoretical and numerical analysis of hyperbolic systems of conservation laws, Master M2 ”mathematical modeling”, 8 lessons of 3h.
- Master: Jing Rebecca Li, Lecturer of course Mathematical and numerical foundations of modeling and simulation using partial differential equations French-Vietnam Master in Applied Mathematics, University of Science, Ho Chi Minh City, 10/2018. 2 weeks.
- Master: P.M. Congedo, Numerical methods in Fluid Mechanics, ENSTA ParisTech, 12 h.
- Doctorat: Houssem Haddar, Lecturer at the kickoff workshop of GDR mecawave. Introduction to Inverse problems (2x1h30).
- Doctorat: P.M. Congedo, Introduction to Uncertainty Quantification, 12h, Doctorate School of University of Bordeaux, France.

10.2.2. Supervision

- Ph.D. A. Bissuel, Linearized Navier Stokes equations for optimization, floating and aeroaccoustic (Dassault Aviation, defended in January 2018). G. Allaire.
- Ph.D. P. Geoffroy, Topology optimization by the homogenization method in the context of additive manufacturing (defended in December 2018). G. Allaire.
- Ph.D. in progress: S. Houbar sur la cavitation dans le fluide caloporteur induite par les mouvements des assemblages d’un réacteur (CEA, to be defended in 2020). G. Allaire and G. Campioni.
- Ph.D. in progress: M. Boissier, Optimisation couplée de la topologie des formes et de la trajectoire de lasage en fabrication additive (to be defended in 2020). G. Allaire and Ch. Tournier.
- L. Rakotondrainibe sur l’optimisation des liaisons enre pièces dans les système mécaniques (Renault, to be defended in 2020). G. Allaire.
- F. Feppon sur l’optimisation topologique de systèmes couplés fluide-solide-thermique (Safran, to be defended in 2020). G. Allaire and Ch. Dapogny.
- Q. Feng sur les éléments finis multi-échelles pour Navier Stokes incompressible en milieu encombré (CEA, to be defended in 2019). G. Allaire and P. Omnes.
- J. Desai sur l’optimisation topologique de structures au comportement non-linéaire avec des méthodes de déformation de maillage (IRT SystemX, to be defended in 2021). G. Allaire and F. Jouve,

- Ph.D. in progress: B. Charfi, Identification of the sigular support of a GIBC, to be defended in 2019, H. Haddar and S. Chaabane
- PhD in progress: K. Napal, Transmission eigenvalues and non destructive testing of concrete like materials , to be defended in 2019, L. Chesnel H. Haddar and L. Audibert
- PhD in progress: M. Kchaou, Higher order homogenization tensors for DMRI modeling, to be defended in 2019, H. Haddar and M. Moakher
- PhD in progress: H. Girardon, Non destructive testing of PWR tubes using eddy current rotating coils, to be defended in 2021, H. Haddar and L. Audibert
- PhD in progress: J. Hao, Thesis topic: Algorithm and software development for analysis and classification of EEG measurements during administration of neuropsychological tests for AD/HD, 2017, J.R. Li and H. Rahioui. PhD stopped in Jan 2019.
- PhD in progress: M. Rihani, Maxwell's equations in presence of metamaterials (to be defended in 2021), A.-S. Bonnet-BenDhia and L. Chesnel.
- PhD in progress: F. Sanson, UQ in systems of solver for the atmospheric reentry (to be defended in July 2019), P.M. Congedo, O. Le Maitre.
- PhD in progress: N. Razaaly, Optimization under uncertainties of ORC turbine cascades, (to be defended in July 2019), P.M. Congedo.
- PhD in progress: M. Rivier, optimization under uncertainty through a Bounding-Box concept (to be defended in May 2020), P.M. Congedo.
- PhD in progress: Joao Reis, Advanced methods for stochastic elliptic PDEs (to be defended in October 2020), P.M. Congedo, O. Le Maitre.
- PhD in progress: G. Gori, Bayesian calibration of complex thermodynamic flows (to be defended in January 2019), P.M. Congedo, O. Le Maitre, A. Guardone.
- PhD in progress: Anabel Del Val, Advanced bayesian methods for aerospace applications (to be defended in October 2020), P.M. Congedo, O. Le Maitre, O. Chazot, T. Magin.
- PhD in progress: J. Carlier, Residual distribution schemes for cavitating two-phase flows (to be defended in October 2019), P.M. Congedo, M. Pelanti, R. Abgrall.
- PhD in progress: P. Novello, Deep learning for reentry atmosperic flows (to be defended in November 2021), P.M. Congedo, D. Lugato, G. Poette.
- PhD in progress: E. Solai, Virtual Prototyping of the EVE expander (to be defended in October 2021), P.M. Congedo, H. Beaugendre.

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

- P.M. Congedo is Deputy Coordinator of "Maths/Engineering" Program of the Labex Mathématiques Hadamard.
- J.R. Li is Member Elu of Inria Commission d'Evaluation, 2015-present.

10.3.2. Internal action

- L. Chesnel provided some numerical experiments used in the exhibition "Rencontres diffractantes : quand les mathématiques inspirent l'art...". This exhibition was presented at Ensta ParisTech and at Inria Saclay.
- P.M. Congedo presented some research activities in aerospace in the context of UnithéauCafé du centre Inria de Saclay Ile-de-France.
- P.M. Congedo made a presentation in the context of the Fete de la science 2018, to several groups of young students (around 10-12 years old).

- H. Haddar made a joint presentation with O. Bunau from Xenocs on nanoparticle imaging using small angle X-ray diffraction technology in the context of UnithéauCafé at Inria de Saclay Ile-de-France.

11. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] L. AUDIBERT, L. A. CHESNEL, H. HADDAR.*Transmission eigenvalues with artificial background for explicit material index identification*, in "Comptes Rendus Mathématique", June 2018, vol. 356, n^{o} 6, p. 626-631 [*DOI :* 10.1016/J.CRMA.2018.04.015], https://hal.archives-ouvertes.fr/hal-01631062

[2] M. BONAZZOLI, F. RAPETTI, C. VENTURINI.*Dispersion analysis of triangle-based Whitney element methods for electromagnetic wave propagation*, in "Applied Mathematics and Computation", February 2018, vol. 319, p. 274-286 [*DOI :* 10.1016/J.AMC.2017.03.026], https://hal.archives-ouvertes.fr/hal-01949026

[3] A.-S. BONNET-BEN DHIA, L. CHESNEL, S. NAZAROV.*Perfect transmission invisibility for waveguides with sound hard walls*, in "Journal de Mathématiques Pures et Appliquées", March 2018, https://hal.archives-ouvertes.fr/hal-01371163

[4] A.-S. BONNET-BEN DHIA, L. CHESNEL, V. PAGNEUX.*Trapped modes and reflectionless modes as eigenfunctions of the same spectral problem*, in "Proceedings of the Royal Society of London. Series A, Mathematical and physical sciences", May 2018, https://hal.archives-ouvertes.fr/hal-01692297

[5] H. BOUJLIDA, H. HADDAR, M. KHENISSI.*The Asymptotic of Transmission Eigenvalues for a Domain with a Thin Coating*, in "SIAM Journal on Applied Mathematics", January 2018, vol. 78, n^{o} 5, p. 2348-2369, https://hal.inria.fr/hal-01646003

[6] F. CAKONI, H. HADDAR, T.-P. NGUYEN.*New interior transmission problem applied to a single Floquet–Bloch mode imaging of local perturbations in periodic media*, in "Inverse Problems", 2018, vol. 35, n^{o} 1, 015009, https://hal.archives-ouvertes.fr/hal-01945600

[7] L. CHESNEL, X. CLAEYS, S. A. NAZAROV.*Oscillating behaviour of the spectrum for a plasmonic problem in a domain with a rounded corner*, in "ESAIM: Mathematical Modelling and Numerical Analysis", September 2018 [*DOI :* 10.1051/M2AN/2016080], https://hal.archives-ouvertes.fr/hal-01240977

[8] L. CHESNEL, S. A. NAZAROV.*Non reflection and perfect reflection via Fano resonance in waveguides*, in "Communications in Mathematical Sciences", 2018, https://hal.archives-ouvertes.fr/hal-01694063

[9] L. A. CHESNEL, S. A. NAZAROV, V. PAGNEUX.*Invisibility and perfect reflectivity in waveguides with finite length branches*, in "SIAM Journal on Applied Mathematics", August 2018, https://hal.archives-ouvertes.fr/hal-01469833

[10] L. CHESNEL, V. PAGNEUX.*Simple examples of perfectly invisible and trapped modes in waveguides*, in "Quarterly Journal of Mechanics and Applied Mathematics", May 2018, https://hal.archives-ouvertes.fr/hal-01593226

[11] A. F. CORTESI, P. M. CONGEDO.*Kriging-sparse Polynomial Dimensional Decomposition surrogate model with adaptive refinement*, in "Journal of Computational Physics", November 2018, https://hal.inria.fr/hal-01914383

[12] H. HADDAR, M. KCHAOU, M. MOAKHER.*The derivation of homogenized diffusion kurtosis models for diffusion MRI*, in "Journal of Magnetic Resonance", 2018, vol. 298, p. 48-57, https://hal.archives-ouvertes.fr/hal-01945614

[13] H. HADDAR, S. MENG.*The spectral analysis of the interior transmission eigenvalue problem for Maxwell's equations*, in "Journal de Mathématiques Pures et Appliquées", December 2018, vol. 120, p. 1-32, https://hal.archives-ouvertes.fr/hal-01945650

[14] V.-D. NGUYEN, J. JANSSON, J. HOFFMAN, J.-R. LI.*A partition of unity finite element method for computational diffusion MRI*, in "Journal of Computational Physics", December 2018, vol. 375, p. 271-290, https://hal.archives-ouvertes.fr/hal-01969367

[15] N. RAZAALY, P. M. CONGEDO.*Novel algorithm using Active Metamodel Learning and Importance Sampling: application to multiple failure regions of low probability*, in "Journal of Computational Physics", April 2018, https://hal.inria.fr/hal-01780050

[16] N. RAZAALY, G. PERSICO, P. M. CONGEDO.*Impact of Geometric, Operational, and Model Uncertainties on the Non-ideal Flow Through a Supersonic ORC Turbine Cascade*, in "Journal of Computational Physics", December 2018, https://hal.inria.fr/hal-01982234

[17] F. SANSON, F. PANERAI, T. E. MAGIN, P. M. CONGEDO.*Robust reconstruction of the catalytic properties of thermal protection materials from sparse high-enthalpy facility experimental data*, in "Experimental Thermal and Fluid Science", March 2018 [*DOI :* 10.1016/J.EXPTHERMFLUSCI.2018.03.028], https://hal.inria.fr/hal-01742595

[18] T. YU, J.-F. CHAIX, L. AUDIBERT, D. KOMATITSCH, V. GARNIER, J.-M. HENAULT.*Simulations of ultrasonic wave propagation in concrete based on a two-dimensional numerical model validated analytically and experimentally*, in "Ultrasonics", February 2019, vol. 92, p. 21 - 34 [*DOI :* 10.1016/J.ULTRAS.2018.07.018], https://hal.archives-ouvertes.fr/hal-01934352

International Conferences with Proceedings

[19] N. RAZAALY, P. M. CONGEDO.*Efficient Computation of Rare Events: Failure Probability and Quantile*, in "MATHIAS-Seminar Total 2018", Serris, France, October 2018, https://hal.inria.fr/hal-01971890

[20] N. RAZAALY, G. GORI, G. IACCARINO, P. M. CONGEDO.*Optimization of an ORC supersonic nozzle under epistemic uncertainties due to turbulence models*, in "GPPS 2019 - Global Power and Propulsion Society", Zurich, Switzerland, January 2019, https://hal.inria.fr/hal-01982227

[21] M. RIVIER, P. M. CONGEDO.*Efficient optimization under uncertainty within the SABBa framework*, in "ECCM-ECFD 2018 - 6th European Conference on Computational Mechanics - 7th European Conference on Computational Fluid Dynamics", Glasgow, United Kingdom, June 2018, https://hal.inria.fr/hal-01841360

[22] M. RIVIER, P. M. CONGEDO.*Low-cost optimization under uncertainty through box representation of robustness measures*, in "Mascot-Num 2018", Nantes, France, March 2018, https://hal.inria.fr/hal-01841352

[23] M. RIVIER, P. M. CONGEDO.*Low-cost optimization under uncertainty through box representation of robustness measures*, in "BIOMA 2018 - 8th International Conference on Bioinspired Optimization Methods and their Applications", Paris, France, May 2018, https://hal.inria.fr/hal-01841372

Conferences without Proceedings

[24] N. RAZAALY, P. M. CONGEDO.*An Efficient Reliability Analysis Tool for the Computation of Low TailProbabilities and Extreme Quantiles in Multiple Failure Regions: Application to Organic RankineCycles*, in "SIAM UQ 2018 - SIAM Conference on Uncertainty Quantification", Garden Grove, California, United States, April 2018, https://hal.inria.fr/hal-01971935

[25] N. RAZAALY, G. PERSICO, P. M. CONGEDO.*Robust Optimization of a Supersonic ORC TurbineCascade under a probabilistic constraint: a Quantile Formulation*, in "NICFD 2018", Bochum, Germany, October 2018, https://hal.inria.fr/hal-01971903

[26] N. RAZAALY, G. PERSICO, G. GORI, P. M. CONGEDO.*Robust Optimization of a Supersonic ORC TurbineCascade: a Quantile-based Approach*, in "MASCOT 2018", Roma, Italy, October 2018, https://hal.inria.fr/hal-01971908

[27] N. RAZAALY, G. PERSICO, G. GORI, P. M. CONGEDO.*Robust Optimization of a Supersonic ORC TurbineCascade: a Quantile-based Approach*, in "ECFD 7 - 7th European Conference on Computational Fluid Dynamic", Glasgow, United Kingdom, June 2018, https://hal.inria.fr/hal-01971926

[28] F. SANSON, C. BERTORELLO, J.-M. BOUILLY, P. M. CONGEDO.*Probabilistic modeling of space object controlled reentry and ground risk estimation*, in "ESA Clean Space Industrial Days", Noordwijk, Netherlands, October 2018, https://hal.archives-ouvertes.fr/hal-01972527

[29] F. SANSON, C. BERTORELLO, C. FINZI, J.-M. BOUILLY, P. M. CONGEDO.*Robust Ground footprint estimation for reentering space objects*, in "4 th International Space Debris Workshop", Darmstadt, Germany, March 2018, https://hal.archives-ouvertes.fr/hal-01972571

[30] F. SANSON, A. EGGELS, D. J. A. CROMMELIN, O. LE MAITRE, P. M. CONGEDO.*Efficient clustering based training set generation for systems of solvers*, in "ECCOMAS Conference", Glasgow, United Kingdom, June 2018, https://hal.archives-ouvertes.fr/hal-01972557

[31] F. SANSON, O. LE MAITRE, P. M. CONGEDO.*Efficient uncertainty propagation in systems of solvers*, in "MATHIAS", Paris, France, October 2018, https://hal.archives-ouvertes.fr/hal-01972538

[32] F. SANSON, O. LE MAITRE, P. M. CONGEDO.*Uncertainty Quantification Framework for Complex Systemsof Solvers*, in "MASCOT2018", Rome, Italy, October 2018, https://hal.archives-ouvertes.fr/hal-01972543

[33] F. SANSON, O. LE MAITRE, P. M. CONGEDO.*Uncertainty Quantification Strategies in Systems of Solvers*, in "SIAM UQ", Los Angeles, United States, April 2018, https://hal.archives-ouvertes.fr/hal-01972566

Scientific Books (or Scientific Book chapters)

[34] L. AUDIBERT, L. CHESNEL, H. HADDAR, K. NAPAL.*Detecting Sound Hard Cracks in Isotropic Inhomogeneities*, in "Advances in Acoustics and Vibration II. ICAV 2018. Applied Condition Monitoring", Springer, September 2019, p. 61-73, https://hal.archives-ouvertes.fr/hal-01945670

Research Reports

[35] M. RIVIER, P. M. CONGEDO.*Surrogate-Assisted Bounding-Box Approach Applied to Constrained Multi-Objective Optimisation Under Uncertainty*, Inria Bordeaux Sud-Ouest ; Inria Saclay, October 2018, n° RR-9214, p. 1-48, https://hal.inria.fr/hal-01897399

[36] M. RIVIER, P. M. CONGEDO.*Surrogate-Assisted Bounding-Box Approach for Optimization Problems with Approximated Objectives*, Inria, February 2018, n° RR-9155, p. 1-35, https://hal.inria.fr/hal-01713043

[37] M. RIVIER, J. LACHAUD, P. M. CONGEDO.*Ablative thermal protection system under uncertainties: effect of pyrolysis gas composition*, Inria Bordeaux Sud-Ouest, May 2018, n° RR-9175, p. 1-23, https://hal.inria.fr/hal-01791072

Other Publications

[38] G. ALLAIRE, P. GEOFFROY-DONDERS, O. PANTZ.*Topology optimization of modulated and oriented periodic microstructures by the homogenization method*, August 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01734709

[39] L. BOURGEOIS, L. CHESNEL, S. FLISS.*On well-posedness of time-harmonic problems in an unbounded strip for a thin plate model*, September 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01883560

[40] F. CAKONI, H. HADDAR, A. LECHLEITER.*On the Factorization Method for a Far Field Inverse Scattering Problem in the Time Domain*, December 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01945665

[41] L. CHESNEL, V. PAGNEUX.*From zero transmission to trapped modes in waveguides*, December 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01952980

[42] F. SANSON, C. BERTORELLO, J.-M. BOUILLY, P. M. CONGEDO.*Breakup prediction under uncertainty: application to Upper Stage controlled reentries from GTO orbit*, October 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01898010

[43] F. SANSON, O. LE MAITRE, P. M. CONGEDO.*Uncertainty Propagation Framework for Systems of Solvers*, July 2018, working paper or preprint, https://hal.inria.fr/hal-01829375

[44] D. WASSERMANN, D. V. NGUYEN, G. GALLARDO, J.-R. LI, W. CAI, V. MENON.*Sensing Von Economo Neurons in the Insula with Multi-shell Diffusion MRI*, 2018, International Society for Magnetic Resonance in Medicine, Poster, https://hal.inria.fr/hal-01807704

Project-Team DISCO

Dynamical Interconnected Systems in COmplex Environments

IN COLLABORATION WITH: Laboratoire des signaux et systèmes (L2S)

IN PARTNERSHIP WITH:
CNRS

CentraleSupélec

RESEARCH CENTER
Saclay - Île-de-France

THEME
Optimization and control of dynamic systems

Table of contents

Project-Team DISCO

Creation of the Team: 2010 January 01, updated into Project-Team: 2012 January 01

Keywords:

Computer Science and Digital Science:

A3.4.4. - Optimization and learning
A3.4.5. - Bayesian methods
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.3. - Discrete Modeling (multi-agent, people centered)
A6.4.1. - Deterministic control
A6.4.3. - Observability and Controlability
A6.4.4. - Stability and Stabilization

Other Research Topics and Application Domains:

B2.2.3. - Cancer
B2.3. - Epidemiology
B3.6. - Ecology
B4.3.3. - Wind energy
B4.4. - Energy delivery
B5.2.3. - Aviation
B7.2.1. - Smart vehicles

1. Team, Visitors, External Collaborators

Research Scientists

Catherine Bonnet [Team leader, Inria, Senior Researcher, HDR]
Frédéric Mazenc [Inria, Senior Researcher from Oct 2018, HDR]
Silviu-Iulian Niculescu [CNRS, Researcher, HDR]

Faculty Members

Sorin Olaru [CentraleSupélec, Professor, HDR]
Guillaume Sandou [CentraleSupélec, Professor, HDR]
Islam Boussaada [IPSA & CNRS, Associate Professor, from Sep 2018, HDR]
Giorgio Valmorbida [Centrale-Supélec, Assistant Professor]

External Collaborators

Islam Boussaada [IPSA & CNRS, until Aug 2018]
Ali Zemouche [Univ de Lorraine]

PhD Students

Saeed Ahmed [Bilkent University & Inria, until Jun 2018]
Caetano Cardeliquio [UNICAMP & Inria]
Jose Hernandez Diez [CNRS & San Luis Potosi,from Jul 2018]
Daniel Ioan [Univ Paris-Sud, from Nov 2017]
Alejandro Martinez Gonzalez [CNRS & San Luis Potosi, from Jul 2018]
Clementine Straub [RTE, from Apr 2018]

Post-Doctoral Fellow

Le Ha Vy Nguyen [Inria, until May 2018]

Visiting Scientists
Stefanella Boatto [Univ Federale Rio de Janeiro, until Mar 2018]
Yutaka Yamamoto [Kyoto University, 6 Sept - 6 Nov 2018]

Administrative Assistant
Katia Evrat [Inria]

2. Overall Objectives

2.1. Objectives

The goal of the project is to better understand and well formalize the effects of complex environments on the dynamics of the interconnections, as well as to develop new methods and techniques for the analysis and control of such systems.

It is well-known that the interconnection of dynamic systems has as consequence an increased complexity of the behavior of the total system.

In a simplified way, as the concept of dynamics is well-understood, the interconnections can be seen as associations (by connections of materials or information flows) of distinct systems to ensure a pooling of the resources with the aim of obtaining a better operation with the constraint of continuity of the service in the event of a fault. In this context, the environment can be seen as a collection of elements, structures or systems, natural or artificial constituting the neighborhood of a given system. The development of interactive games through communication networks, control from distance (e.g. remote surgical operations) or in hostile environment (e.g. robots, drones), as well as the current trend of large scale integration of distribution (and/or transport and/or decision) and open information systems with systems of production, lead to new modeling schemes in problems where the dynamics of the environment have to be taken into account.

In order to tackle the control problems arising in the above examples, the team investigates new theoretical methods, develops new algorithms and implementations dedicated to these techniques.

3. Research Program

3.1. Analysis of interconnected systems

The major questions considered are those of the characterization of the stability (also including the problems of sensitivity compared to the variations of the parameters) and the determination of stabilizing controllers of interconnected dynamic systems. In many situations, the dynamics of the interconnections can be naturally modelled by systems with delays (constant, distributed or time-varying delays) possibly of fractional order. In other cases, partial differential equations (PDE) models can be better represented or approximated by using systems with delays. Our expertise on this subject, on both time and frequency domain methods, allows us to challenge difficult problems (e.g. systems with an infinite number of unstable poles).

- Robust stability of linear systems

  Within an interconnection context, lots of phenomena are modelled directly or after an approximation by delay systems. These systems may have constant delays, time-varying delays, distributed delays ...

  For various infinite-dimensional systems, particularly delay and fractional systems, input-output and time-domain methods are jointly developed in the team to characterize stability. This research is developed at four levels: analytic approaches (H_∞-stability, BIBO-stablity, robust stability, robustness metrics) [1], [2], [5], [6], symbolic computation approaches (SOS methods are used for determining easy-to-check conditions which guarantee that the poles of a given linear system are not in the closed right half-plane, certified CAD techniques), numerical approaches (root-loci, continuation methods) and by means of softwares developed in the team [5], [6].

- Robustness/fragility of biological systems

  Deterministic biological models describing, for instance, species interactions, are frequently composed of equations with important disturbances and poorly known parameters. To evaluate the impact of the uncertainties, we use the techniques of designing of global strict Lyapunov functions or functional developed in the team.

  However, for other biological systems, the notion of robustness may be different and this question is still in its infancy (see, e.g. [65]). Unlike engineering problems where a major issue is to maintain stability in the presence of disturbances, a main issue here is to maintain the system response in the presence of disturbances. For instance, a biological network is required to keep its functioning in case of a failure of one of the nodes in the network. The team, which has a strong expertise in robustness for engineering problems, aims at contributing at the develpment of new robustness metrics in this biological context.

3.2. Stabilization of interconnected systems

- Linear systems: Analytic and algebraic approaches are considered for infinite-dimensional linear systems studied within the input-output framework.

  In the recent years, the Youla-Kučera parametrization (which gives the set of all stabilizing controllers of a system in terms of its coprime factorizations) has been the cornerstone of the success of the H_∞-control since this parametrization allows one to rewrite the problem of finding the optimal stabilizing controllers for a certain norm such as H_∞ or H_2 as affine, and thus, convex problem.

  A central issue studied in the team is the computation of such factorizations for a given infinite-dimensional linear system as well as establishing the links between stabilizability of a system for a certain norm and the existence of coprime factorizations for this system. These questions are fundamental for robust stabilization problems [1], [2].

  We also consider simultaneous stabilization since it plays an important role in the study of reliable stabilization, i.e. in the design of controllers which stabilize a finite family of plants describing a system during normal operating conditions and various failed modes (e.g. loss of sensors or actuators, changes in operating points). Moreover, we investigate strongly stabilizable systems, namely systems which can be stabilized by stable controllers, since they have a good ability to track reference inputs and, in practice, engineers are reluctant to use unstable controllers especially when the system is stable.

- Nonlinear systems

  The project aims at developing robust stabilization theory and methods for important classes of nonlinear systems that ensure good controllerperformance under uncertainty and time delays. The main techniques include techniques called backstepping and forwarding, contructions of strict Lyapunov functions through so-called "strictification" approaches [3] and construction of Lyapunov-Krasovskii functionals [4], [5], [6].

- Predictive control

  For highly complex systems described in the time-domain and which are submitted to constraints, predictive control seems to be well-adapted. This model based control method (MPC: Model Predictive Control) is founded on the determination of an optimal control sequence over a receding horizon. Due to its formulation in the time-domain, it is an effective tool for handling constraints and uncertainties which can be explicitly taken into account in the synthesis procedure [7]. The team considers how mutiparametric optimization can help to reduce the computational load of this method, allowing its effective use on real world constrained problems.

  The team also investigates stochastic optimization methods such as genetic algorithm, particle swarm optimization or ant colony [8] as they can be used to optimize any criterion and constraint whatever their mathematical structure is. The developed methodologies can be used by non specialists.

3.3. Synthesis of reduced complexity controllers

- PID controllers

  Even though the synthesis of control laws of a given complexity is not a new problem, it is still open, even for finite-dimensional linear systems. Our purpose is to search for good families of "simple" (e.g. low order) controllers for infinite-dimensional dynamical systems. Within our approach, PID candidates are first considered in the team [2], [67].
- Predictive control

  The synthesis of predictive control laws is concerned with the solution of multiparametric optimization problems. Reduced order controller constraints can be viewed as non convex constraints in the synthesis procedure. Such constraints can be taken into account with stochastic algorithms.

Finally, the development of algorithms based on both symbolic computation and numerical methods, and their implementations in dedicated Scilab/Matlab/Maple toolboxes are important issues in the project.

4. Application Domains

4.1. Analysis and Control of life sciences systems

The team is involved in life sciences applications. The two main lines are the analysis of bioreactors models and the modeling of cell dynamics in Acute Myeloblastic Leukemias (AML) in collaboration with St Antoine Hospital in Paris. A recent new subject is the modelling of Dengue epidemics.

4.2. Energy Management

The team is interested in Energy management and considers optimization and control problems in energy networks.

5. New Software and Platforms

5.1. FEMMES

KEYWORD: Linear system
FUNCTIONAL DESCRIPTION: The Software FEMMES makes it possible to perform simulations for observers which converge to the exact value of the solutions of a studied system in finite time. The considered systems are linear continuous-time time-invariant systems.

- Partner: Inria
- Contact: Frédéric Mazenc

6. New Results

6.1. Spectral abscissa characterization for Time-delay systems

Participants: Islam Boussaada, Silviu-Iulian Niculescu, Sami Tliba [Université Paris Sud], Thomas Vyhlidal [Czech technical university in Prague], Karim Trabelsi [IPSA].

It is well known that the spectral abscissa of a given dynamical is nothing but the corresponding solutions' exponential decay. The analytical characterization of the spectral abscissa for infinite dimensional dynamical systems is an old problem which is still nowadays a question of ongoing interest due to its links with stability problems. We produced several works in this topic dealing with reduced order retarded Time-delay systems and emphasized a property that we call *multiplicity induced-dominancy*. In the paper [13], the interest of using time-delay in the controller design as a control parameter is underlined and the way to assign a dominant spectral value is demonstrated. As a matter of fact, it is shown that the multiplicity of given spectral value may reach the degree of the corresponding quasipolynomial. Furthermore, when this holds, then using a particular factorization of the quasipolynomial, such a multiple spectral value is shown to be the corresponding spectral abscissa. A generalization of such a result to generic second order retarded equation with a single delay is established in [12]. More precisely, a parametric characterization of the spectral abscissa is established using the principle argument theorem. Furthermore, in the work [45], the potential applicability of such a parametric characterization in controller design in concrete application is demonstrated. As a matter of fact, a third order retarded system modeling the dynamics of Mach number in a wind tunnel is considered and a delayed controller design based on the spectral abscissa assignment is proposed.

6.2. Poles placement for reduced order Time-delay systems

Participants: Souad Amrane [University Mouloud Mammeri], Islam Boussaada, Fazia Bedouhen [University Mouloud Mammeri], Silviu-Iulian Niculescu, Matej Kure [Czech technical university in Prague], Wim Michiels [KU Leuven], Thomas Vyhlidal [Czech technical university in Prague].

It is well known in dynamical system theory that real spectral values correspond to non oscillating solutions. In the paper [11]we made a connexion between the degree of a given quasipolynomial and the admissible number of non oscillating modes for the corresponding Time-delay system. More precisely, we have shown that the assignment of at most n real spectral values is possible for generic quasipolynomial function of degree n. Namely, explicit formulas on the quasipolynomial's coefficients guaranteeing the coexistence of n negative spectral values are obtained. Furthermore, a new quasipolynomial factorization technique, analogous to the one we developed for multiple spectral values for the proof of the dominancy of n distinct negative spectral values is obtained.

In the paper [23] a robust alternative of the delayed resonator is proposed by spectral approach where a double root assignment at the excitation frequency is proposed. Such an excitation frequency is projected to widening the stop-band in the active absorber frequency response. It is shown that the performance sensitivity to the mismatch between the design and true excitation frequency is considerably decreased. Additionally, the overall scheme is supplemented by a control loop which improves the stability margin.

6.3. Asymptotic behavior of critical imaginary roots for retarded differential equations

Participants: Islam Boussaada, Jie Chen [City University of Hong Kong], Liana Felix [Universidad Autonoma de San Luis Potosi], Keqin Gu [Southern Illinois University], Fernando Mendez-Barrios [Universidad Autonoma de San Luis Potosi], Dina Irofti, Silviu-Iulian Niculescu, Alejandro Martinez.

The behavior of characteristic roots of time-delay systems, when the delay is subject to small variations is investigated in [26]. We performed an analysis by means of the Weierstrass polynomial which are employed to study the stability behavior of the characteristic roots with respect to small variations on parameters. Analytic description and splitting properties of the Puiseux series expansions of critical roots are characterized by allowing a full description covering all the cases that can be encountered.

In the paper [21] the migration of double imaginary roots of the systems characteristic equation when two parameters are subjected to small deviations is geometrically investigated. Under the least degeneracy assumptions, the local stability crossing curve is shown to have a cusp at the point that corresponds to the double root, which divides the neighborhood of this point into two sectors (called S-sector and a G-sector). We have shown that when the parameters move into the G-sector, one of the roots moves to the right half-plane, and the other moves to the left half-plane. However, when the parameters move into the S-sector, both roots move either to the left half-plane or the right half-plane depending on the sign of a quantity that depends on the characteristic function and its derivatives up to the third order.

6.4. Stability analysis of retarded differential equations with delay-dependent coefficients

Participants: Islam Boussaada, Silviu-Iulian Niculescu, Chi Jin [IPSA], Keqin Gu [Southern Illinois University].

Retarded dynamical systems with delay dependent coefficients is a class of systems which is frequently encountered in various scientific and engineering applications. The paper [36] provides an overview of the stability analysis of such systems which generalizes those on systems with delay-independent coefficients. Methods of analysis for systems with a single delay and commensurate delays are presented, their application to output feedback control and a geometric perspective that establishes a link between systems with and without delay-dependent coefficients.

The paper [22] presents a systematic method to analyse the stability of systems with single delay in which the coefficient polynomials of the characteristic equation depend on the delay. With respect to the literature on the topic, a less restrictive method to analyse stability is presented. It is found that a much richer behavior is possible when the restrictive assumptions are removed. The interval of interest for the delay is partitioned into subintervals so that the magnitude condition generates a fixed number of frequencies as functions of the delay within each subinterval. The crossing conditions are expressed in a general form, and a simplified derivation for the first-order derivative criterion is obtained.

6.5. Stability and Stabilisability Through Envelopes for Retarded and Neutral Time-Delay Systems

Participants: Catherine Bonnet, Caetano Cardeliquio, Silviu Niculescu, André Fioravanti [FEM-UNICAMP, Brazil].

Through an LMI approach it was possible to determine envelopes and use them not only to study stability but to design robust controllers for retarded and neutral time-delay systems. The controller designed is robust to parametric uncertainties and can guarantee delay independant stability or delay-dependant $\alpha - stability$ [46].

6.6. Some remarks on the Walton and Marshall method for neutral delay systems

Participants: Catherine Bonnet, Islam Boussaada, Le Ha Vy Nguyen, Marianne Souaiby.

The Walton and Marshall method allows to determine stability windows of delay systems of the retarded and neutral type. We noticed that some delay systems of the neutral type do not behave as claimed in [66] and analyzed carefully the position of the poles of such systems in the right half-plane.

6.7. Local Analysis of Lurie Systems

Participants: Elena Panteley [L2S,CNRS], Stephen Duncan [University of Oxford], Thomas Lathuiliere [University of Oxford], Giorgio Valmorbida.

An important aspect of nonlinear systems is the fact that stability might only be a local property. This means that associated to a stable equilibrium point or periodic trajectory, there is a region of attraction. Such a region is formed by points of trajectories converging to the stable sets. An important task of practical interest is then to estimate these regions via numerical methods that rely on the model of the system. As an illustration, it might be of interest to know the region of safe operation of an electric motor in order to preserve its integrity or, in the case of an autonomous vehicle, limit the operating condition for safety purposes.

For the particular class of Lurie systems, namely systems defined by the interconnection of a linear system and a static nonlinearity, it is possible to compute estimates based on sector inequalities characterizing the nonlinearities in the system. If further information, such as the slope of the nonlinearity is available, one can better characterize local properties such as regions of stability, and input-output relations such as reachable sets and local nonlinear gains.

To obtain these characterizations we rely on numerical methods based on convex optimization. These methods are based on the solution of Lyapunov inequalities yielding Lyapunov functions that are quadratic on both the states and the nonlinearity and has an integral term on the nonlinearity [39].

Moreover, whenever a more precise characterization of the nonlinearity is at hand as for instance nonlinearities having rational Jacobian, one can generalize the local analysis methods using polynomial optimization. This includes the case of standard Lurie systems by considering the interconnection of a polynomial system with static sector nonlinearities that have rational Jacobian. In this setting we have proposed conditions that relax the requirement on the candidate Lyapunov function [17], which serve as stability certificates, from being sum-of-squares of polynomial with respect to the nonlinearities and the Lurie-Postnikov terms from being non-negative.

Further to the stability analysis we were interested in another important phenomenon and its analysis through numerical methods : the existence of limit cycles on nonlinear systems. Such a phenomenon is relevant since it can be used as a method to design stable oscillators with known amplitude and frequency but also to evaluate and suppress undesirable oscillations in engineered systems. In order to proceed with this analysis we have limited our attention to a particular class of systems defined by a Liénard systems and formulate sufficient conditions for existence and uniqueness of limit cycles for systems with a non-differentiable vector field. As an application we consider the example of a linear system with saturation [24]. Moreover, for planar saturating systems we present sufficient conditions for the existence of periodic orbits and we characterize inner and outer sets bounding the periodic orbits. A method to build these bounds, based on the solution to a convex optimization problem is proposed and numerical examples optimizing the region bounding the limit cycle illustrate the technique [25].

6.8. New advances on backstepping

Participants: Frederic Mazenc, Michael Malisoff [LSU], Laurent Burlion [ONERA Toulouse], Jerome Weston [LSU].

We worked on the problem of improving a fundamental control design technique for nonlinear systems called backstepping by using a fundamentally new approach which consists in introducing in the control artificial delays or using dynamic extensions.

In [28], we provided backstepping results for a large class of partially linear systems with an arbitrarily large number of integrators. We proposed control laws whose size respects some constaints given a priori. The key tool is a dynamic extension that contains only one artificial delay, which is in sharp contrast with our prior contributions. We also showed that the closed-loop system is robust, in the input-to-state stability sense, with respect to a large class of model uncertainties, and robust with respect to delays in the measurements. We illustrated the result using an example that is beyond the scope of classical backstepping.

The paper [57] also provides a crucial backstepping result. We explained how globally asymptotically stabilizing output feedbacks can be constructed for a family of nonlinear systems using only a dynamic extension and a "Converging Input-Converging State" assumption and no additional delays. The technique presents several advantages. It provides control laws whose expressions are simple. It makes it possible to stabilize systems in the presence of uncertain terms, which are not necessarily of class C^1 and which prevent the use of the classical backstepping technique. It applies in cases where only part of the state variables can be measured.

6.9. Time-varying systems with delay and Switched Systems

Participants: Frederic Mazenc, Michael Malisoff [LSU], Saeed Ahmed [Inria], Hitay Ozbay [Blikent University, Turkey].

The family of the switched systems is frequently encountered in practice. It can be used to approximate time-varying systems to ease their stability analysis or control.

In [29] we provided theoretical results for the stability and robustness analysis of nonlinear switched time-varying systems with uncertainties and time-varying delays. The delays are allowed to be discontinuous and arbitrarily long with known upper bounds. We established the results via an adaptation of Halanay's inequality and a trajectory based technique. Also, we used the results for designing switched controllers that stabilize linear time-varying systems with time-varying delays.

The contribution [10] proposed a new technique of construction of observers making possible to stabilize by output feedback a class of continuous-time switched linear systems with a time-varying delay in the output. The motivation of this paper is strong: frequently measurements are affected by pointwise time-varying delays. For stability analysis, we developped an extension of the trajectory based approach. A stability condition is given in terms of the upper bound on the time-varying delay to ensure global exponential stability of the switched feedback systems. It is worth observing that the main result applies in cases where some of the subsystems of the switched system are not stabilizable and not detectable.

The paper [27] is also devoted to classes of nonlinear time-varying continuous-time systems with outputs. For a first family of systems, we builded an observer in the case where a state dependent disturbance affects the linear approximation. A fundamental feature of our observer is the fact that it converges after a predetermined finite time. When the disturbances are the zero functions, it provides exact values of the state and it provides an approximate estimate when there are nonzero disturbances. We used this construction to design a globally exponentially stabilizing dynamic output feedback for a second family of nonlinear systems whose outputs are only available on some finite time intervals. Our technique consists in switching between control laws. We applied the control design to the controlled Mathieu equation, which arises in the study of vibrations of an elliptic membrane.

The paper [38] is devoted to a stability analysis for a class of nonlinear systems with a time-varying delay taking both large and small values in an alternating manner, precluding the application of most of the classical control design techniques. The type of assumption we imposed is the following: we imposed on the delay to be "small" on "long" time intervals and possibly "large" on "small" time-intervals. Bearing in mind this key property, we first introduced the concept of delay-hybrid-dependent stability, which grasps the features of the delays described above and represented the studied system as a system with a switched delay. Then by using switching techniques and Lyapunov-Krasovskii functionals (LKFs), we provided a new stability criterion.

6.10. Observers

Participants: Frederic Mazenc, Michael Malisoff [LSU], Saeed Ahmed [Inria], Ali Zemouche [CRAN], Rajesh Rajamani [University of Minneapolis, USA], Maruthi Akella [University of Texas, USA].

We produced several works which pertain to the case where only a part of the state variables can be measured.

In the paper [58], we adopted a technique based on the indroduction of several observers in cascade (such a cascade is called 'sequential observer') for a class of time-varying linear systems in which the inputs and outputs containing sampling and arbitrarily long delays. The observers are of a continuous-discrete type. We used the observers to design controllers that ensure a strong robustness property with respect to uncertainties in the system and the output, under delays and sampling. A fundamental aspect of the approach is that it produces the observers and controllers without distributed terms. We have assessed the performance of the control laws through two examples, which inlcude a DC motor model that illustrates the utility of the work in engineering applications.

In two papers, we developped the theory of the finite time observers. In [53], we study a class of linear continuous-time time-varying systems with piecewise continuous disturbances and piecewise constant outputs. Under a classical assumption of observability, we designed a new type of observers to estimate the solutions of the system in a predetermined finite time. In contrast to the well-established finite time observer design techniques which estimate the system state using a continuous output, our proposed observer applies when only piecewise constant measurements are available. In [54], we construct finite-time reduced order observers for a broad family of nonlinear time-varying continuous-time systems. The motivation for this is the fact that in practice the time-varying aspect of a system may be an obstacle to the design of full-order finite-time observers, but not for the design of reduced order ones. We illustrated our results using a tracking problem for nonholonomic systems in chained form.

Two of our works present construction of asymptotic observers without delay. The paper [61] solves an H_∞ observer design problem for a class of descriptor nonlinear systems. The method we established is theoretical and can be applied to many automatic control design problems such as unknown input estimation problem, which plays an important role in control systems, namely for diagnosis and fault tolerant control. The design relies on the Linear Matrix Inequality condition (LMI) technique. We applied our resut to a model of a flexible joint robot system.

The work [16] is dedicated to the design of a smooth six-degree-of-freedom observer to estimate the incorporating linear and angular velocity, called dual angular velocity, for a rigid body. The approach is based on the dual-quaternion description and we proved that the estimation errors exhibit asymptotic convergence. Furthermore, to achieve tracking control objective, we combined the proposed observer with an independently designed proportional-derivative-like feedback control law (using full-state feedback), and a special Lyapunov "strictification" process is employed to ensure a separation property between the observer and the controller. We performed numerical simulations for a prototypical spacecraft hovering mission application.

6.11. Stabilization of various systems with pointwise delays

Participants: Frederic Mazenc, Michael Malisoff [LSU], Delphine Bresch-Pietri [Mines Paris Tech.], Nicolas Petit [Mines Paris Tech.], Robledo Gonzalo [Univ. de Chile, Chile], Maruthi Akella [University of Texas, USA], Xi-Ming Sun [Dalian University of Technology, China], Xue-Fan Wang [Dalian University of Technology, China].

The presence of delays to big for being neglected is an obstacle to the design of stabilizing controllers in many cases. We have made efforts to overcome this challenge by developping several techniques.

In the paper [14], we investigated the design of a prediction-based controller for a linear system subject to a problematic time-varying input delay: the delay we considered is not necessarily "First-In/First-Out". The feedback law we proposed uses the current delay value in the prediction. It does not exactly compensate the delay in the closed-loop dynamics but does not require to predict future delay values, contrary to classical prediction techniques. Modeling the input delay as a transport Partial Differential Equation, we proved asymptotic stabilization of the system state, provided that the average L_2-norm of the first derivative of the delay over some time-window is sufficiently small and that the average time between two discontinuities (average dwell time) is sufficiently large.

In the paper [51], we adopted another type of strategy: we used a new sequential predictors approach to build uniformly globally exponentially stabilizing feedback controls for a large class of linear time-varying systems that contain an arbitrary number of different delays. This allows different delays in different components of the input. We illustrated our work in an example from identification theory, and in an Euler-Lagrange system arising from two-link manipulator systems.

The paper [31] continues our works on the chemostat model with an arbitrary number of competing species, one substrate, and constant dilution rates. We allowed delays in the growth rates and additive uncertainties. Using constant inputs of certain species, we derived bounds on the sizes of the delays that ensure asymptotic stability of an equilibrium when the uncertainties are zero, which can allow persistence of multiple species. In the presence of delays and uncertainties, we provided bounds on the delays and on the uncertainties that ensure, with respect to uncertainties, the robustness property called "input-to-state stability".

6.12. Low complexity constrained control using higher degree Lyapunov functions

Participants: Sarmad Munir [NTNU, Trondheim], Sorin Olaru, Morten Hovd [NTNU, Trondheim].

Explicit Model Predictive Control often has a complex solution in terms of the number of regions required to define the solution and the corresponding memory requirement to represent the solution in the online implementation. An alternative approach to constrained control is based on the use of controlled contractive sets. However, polytopic controlled contractive sets may themselves be relatively complex, leading to a complex explicit solution, and the polytopic structure can limit the size of the controlled contractive set. Our recent results [33] develop a method to obtain a larger controlled contractive set by allowing higher order functions in the definition of the contractive set, and explores the use of such higer-order contractive sets in controller design leading to a low complexity explicit control formulation.

6.13. Characterization of ultimate bounds for systems with state-dependent disturbances

Participants: Sorin Olaru, Hiroshi Ito [Kyushu Institute of Technology, Japan].

The work [37] pursues a framework of set characterization of dynamical systems with state-dependent disturbances. It aims to propose a new approach to analysis and design of nonlinear systems involving non-differentiability and asymmetric components which hamper application and effectiveness of local linearization methods. Several characterizations of ultimate bounds are developed. The utility of shifting the fix point is formulated as a parametrization of the ultimate bounds.

6.14. Combinatorial Approach towards Multi-Parametric Quadratic Programming based on Characterizing Adjacent Critical Regions

Participants: Parisa Ahmadi-Moshkenani [NTNU, Trondheim], Sorin Olaru, Tor Johansen [NTNU, Trondheim].

Several optimization-based control design techniques can be cast in the form of parametric optimization problems. The multi-parametric quadratic programming (mpQP) represents a popular class often related to the control of constrained linear systems. The complete solution to mpQP takes the form of explicit feedback functions with a piecewise affine structure, valid in polyhedral partitions of the feasible parameter space known as critical regions. The recently proposed combinatorial approach for solving mpQP has shown better efficiency than geometric approaches in finding the complete solution to problems with high dimensions of the parameter vectors. The drawback of this method, on the other hand, is that it tends to become very slow as the number of constraints increases in the problem. This work [9] presents an alternative method for enumerating all optimal active sets in a mpQP based on theoretical properties of adjacent critical regions and their corresponding optimal active sets. Consequently, it results in excluding a noticeable number of feasible but not optimal candidate active sets from investigation. Therefore, the number of linear programs that should be solved decreases noticeably and the algorithm becomes faster. Simulation results confirm the reliability of the suggested method in finding the complete solution to the mpQPs while decreasing the computational time compared favourably with the best alternative approaches.

6.15. Active vibration damping in a mechanical structures

Participants: Islam Boussaada, Silviu-Iulian Niculescu, Sami Tliba [Université Paris Sud], Thomas Vyhlidal [Czech technical university in Prague], Daniela Danciu [University of Craiova].

In the work [13], an aluminium-based flexible structure embedded in a mobile support subjected to an acceleration is considered. Such a flexible beam is equipped with two piezoelectric patches. One of them is used as an actuator and the second acts as a sensor. These patches are supposed to be rigidly bounded on the beam, one on each side, located at the clamped edge. The whole device is called a piezo-actuated beam which is generally modeled by Euler-Bernoulli equations. Finite element modeling is then applied to reduce the PDE system to a linear finite-dimensional system. Then, the peak of resonance of the first bending mode is damped by using a delayed output-feedback controller, without affecting the neglected vibrating modes in the reduced order model. The proposed controller design is based on the spectral abscissa characterization using the multiplicity property.

6.16. Landing of a civil aircraft

Participants: Frederic Mazenc, Michael Malisoff [LSU], Laurent Burlion [ONERA Toulouse], Victor Gibert [Airbus Toulouse].

In this work and the following, we applied the technique of [28] to problems arising from applications. The paper [56] is devoted to the problem of stabilizing a nonlinear system approximated in a neighborhood of the origin by a saturated chain of integrators when the variables are not accurately measured. We used our control design to solve a control problem that arises in the context of vision based landing of a civil aircraft. In [55], we solved the problem of stabilizing a nonlinear system when the variables are not accurately measured and cannot be differentiated. The proposed method was first motivated and thus finally applied to the vision based control problem of a landing airliner.

6.17. Power electronics devices

Participants: Frederic Mazenc, Alessio Iovine [Efficacity, France].

The contribution [50] is distinct from the papers mentioned above because it uses more traditional backstepping tools. It is devoted to power electronics devices. We proposed a nonlinear control law for a DC/DC boost converter dedicated to extract the maximum power from a photovoltaic (PV) array, taking into account the constraints of the control action. We performed simulations on SimPowerSystems to validate how the developed control strategy is able to properly control the converter.

6.18. Wind Farm Distributed PSO-based Control for Constrained Power Generation Maximization

Participants: Nicolo Gionfra [L2S], Guillaume Sandou, Houria Siguerdidjane [L2S], Damien Faille [EDF], Philippe Loevenbruck [EDF].

A novel distributed approach to treat the wind farm (WF) power maximization problem accounting for the wake interaction among the wind turbines (WTs) is presented. Power constraints are also considered within the optimization problem. These are either the WTs nominal power or a maximum allowed power injection, typically imposed by the grid operator. The approach is model-based. Coupled with a distributed architecture it allows fast convergence to a solution, which makes it exploitable for real-time operations. The WF optimization problem is solved in a cooperative way among the WTs by introducing a new distributed particle swarm optimization algorithm, based on cooperative co-evolution techniques. The algorithm is first analyzed for the unconstrained case, where we show how the WF problem can be distributed by exploiting the knowledge of the aerodynamic couplings among the WTs. The algorithm is extended to the constrained case employing Deb's rule. Simulations are carried out on different WFs and wind conditions, showing good power gains and fast convergence of the algorithm. To appear in Rnewable Energy, 2019.

6.19. Wind Farm Distributed PSO-based Control for Constrained Power Generation Maximization

Participants: Sophie Frasnedo [Safran Electronics and Defense], Guillaume Sandou, Gilles Duc [L2S], Philippe Feyel [Safran Electronics and Defense], Cedric Chapuis [Safran Electronics and Defense].

The inertial stabilisation of the line of sight of an imager fixed on a mobile carrier is considered in order to acquire good quality images despite the disturbances generated by the carrier.

A double stage mechanical stabilisation architecture is proposed, where a second stabilisation stage, based on a piezoelectric actuator, is added to the usual structure. The piezoelectric actuator transfer function and hysteresis are characterized through experiments.

In order to design the controllers of both stages, a high-level image quality criterion (the Modulation Transfer Function MTF) is considered, together with design constraints on the main variables of interest. The criterion and the constraints are evaluated by realistic simulations based on some input and noise profiles measured on a real-life system. The MTF evaluation being time-consuming, a Bayesian optimisation method specially dedicated to expensive-to-evaluate functions is used to obtain the parameters of the controllers. The obtain experimental results are displayed and their performances discussed. To appear in the International Journal of Systems and Sciences in 2019.

6.20. Model Identification for Demand-Side Management of District Heating Substations

Participants: Nadine Aoun [L2S, CEA-LITEN, ADEME], Roland Baviere [CEA-LITEN], Mathieu Vallee [CEA-LITEN], Guillaume Sandou.

Demand-Side Management (DSM) strategies, such as load shifting and nighttime set-back, exploit the thermal inertia of buildings to make the operation of District Heating Systems (DHSs) more efficient. The control strategy requires a building model to assess the flexibility of buildings in handling demand modulation, without jeopardizing the thermal comfort. Reduced Order Models (ROMs) with few parameters are often used for this end; in many previous works their parameters have been identified using time-series data including indoor temperature measurements. However, at a city scale and due to privacy rights, such internal signals are usually unavailable. Thereby, identifying the ROM shall rely solely on measurements available at the substation level.

In our work, we develop and demonstrate a method respecting this practical constraint to identify a first and a second order building model. In literature, a rather simplified approach had been proposed to derive a first order building model from substation measurements. We compare the performance of our methodology with respect to the latter, using the same model structure. As for the second order model, its structure is more relevant to account for different dynamics in buildings equipped with hydronic heating systems or featuring important internal thermal inertia. Data used for the identification is restricted to the heat flux delivered from the DHS, both supply and return water temperatures, mass flowrate across the substation's heat-exchangers and the outdoor temperature. Validation of the proposed approach is carried out using a representative white-box model of a building and its substation written in the Modelica language. Implementation of advanced control strategies for DHSs based on this model identification is in prospect.

6.21. Mathematical Modelling of Acute Myeloid Leukemia

Participants: Catherine Bonnet, Jean Clairambault [MAMBA project-team], François Delhommeau [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Walid Djema, Emilia Fridman [Tel-Aviv University], Pierre Hirsch [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Frédéric Mazenc, Hitay Özbay [Bilkent University].

Our project is about the modeling and analysis of healthy and unhealthy cell population dynamics, with a particular focus on hematopoiesis, which is the process of blood cell production and continuous replenishment. We point out that medical research is now looking for new combined targeted therapies able to overcome the challenge of cancer cells (e.g. to stop overproliferation, to restore normal apoptosis rates and differentiation of immature cells, and to avoid the high toxicity effects that characterize heavy non-selective chemotherapy). In that quest, the ultimate goal behind mathematical studies is to provide some inputs that should help biologists to suggest and test new treatment, and to contribute within multi-disciplinary groups in the opening of new perspectives against cancer. Thus, our research project is imbued within a similar spirit and fits the expectations of a better understanding of the behavior of healthy and unhealthy blood cell dynamics. It involve intensive collaboration with hematologists from Saint Antoine hospital in Paris, and aims to analyze the cell fate evolution in treated or untreated leukemia, allowing for the suggestion of new anti-leukemic combined chemotherapy.

Cells have amazing features that allow them to guide their development paths and determine their individual and collective fates. Dedifferentiation and transdifferentiation (cell plasticity) are little understood phenomena that allow cells to regress from an advanced differentiated state to a less differentiated one, including the case where cells lose their specific function and become stem cells.

We have introduced cell plasticity into a class of mathematical models we are interested in. We explored a new model involving a dedifferentiation function in the case of two cell maturity stages (stem cells and progeny). We have highlighted the role that dedifferentiation may have in the survival of cancer cells during therapy. The latter hypothesis appears to be in line with some medical observations [48].

We have also developed and analyzed a model taking into the account the fact that few cells of the proliferating compartment may be arrested during an unlimited time [49].

6.22. Analysis of Dengue Fever SIR Model with time-varying parameters

Participants: Stefanella Boatto [Univ Feder Rio de Janeiro], Catherine Bonnet, Frédéric Mazenc, Le Ha Vy Nguyen.

Migratory fluxes of humains and of insects of various species have favoured the spreading of diseases worldwide. In particular the Ae Aegypti and Ae Albopictus mosquitoes of the Aedes family are vectors able to transmit and spead among humans a variety of diseases: Dengue, Zika, Chikungunya, Yellow fever and the newly discovered Mayaro.

We have continued to analyze SIR models with time-varying parameters to predict dengue epidemics and compared numerical simulations with real data from Dengue epidemics in Rio de Janeiro in order to estimate the infectivity rate and predict what are the periods more at risk of infection [63], [41].

7. Partnerships and Cooperations

7.1. National Initiatives

7.1.1. Industrial-Academic Institute

Guillaume Sandou is the head of the RISEGrid Institute. The Institute is dedicated to the study, modelling and simulation of smart electric distribution grids and their interactions with the whole electric power system. It is located in CentraleSupélec and gathers about 20 people (academic and industrial researchers, PhD students, post-doctoral researchers). The Institute has been renewed in 2018 for 5 new years.

7.2. European Initiatives

7.2.1. Collaborations in European Programs, Except FP7 & H2020

Program: **COST Action**

Project acronym: FRACTAL

Project title: Fractional-order systems; analysis, synthesis and their importance for future design

Duration: November 2016 - October 2020

Coordinator: Jaroslav Koton Czech Republic

Abstract: Fractional-order systems have lately been attracting significant attention and gaining more acceptance as generalization to classical integer-order systems. Mathematical basics of fractional-order calculus were laid nearly 300 years ago and since that it has gained deeply rooted mathematical concepts. Today, it is known that many real dynamic systems cannot be described by a system of simple differential equation or of integer-order system. In practice we can encounter such systems in electronics, signal processing, thermodynamics, biology, medicine, control theory, etc. The Action will favor scientific advancement in above mentioned areas by coordinating activities of academic research groups towards an efficient deployment of fractal theory to industry applications.

Program: **PHC BALATON**

Project acronym: SadHuB

Project title: Analysis of stabilizability of delayed dynamical system as function of the systems parameters and the time delays with applications to human balancing

Duration: 01/2018-12/2019

Coordinator: Islam Boussaada

Other partners: Budapest University of Technology and Economics, Hungary

Abstract: Motivated by a class of Time-delay systems occurring in modeling of many mechanical engineering applications, this project aims to associate researchers from control theory, applied mathematics and mechanical engineering to build together a general methodology for the analysis and control of mechanical/bio-mechanical structures. In particular, the human balance is often considered as a control system which operates in the presence of delays, primarily due to the time it takes to acquire the information needed for decision-making, to create control decisions, and to execute these decisions. A particular interest will be devoted to the delayed human balance, where a depthful study of the delay effect on the stability is expected.

Program: **PHC BRANCUSI**

Project acronym: Proco

Project title: systems with propagation: new approaches in control design for oscillations quenching

Duration: 01/2017-12/2018

Coordinator: Islam Boussaada

Other partners: Craiova University, Romania

Abstract: Propagation systems are met and analyzed in various fields, in natural sciences (Physics, Chemistry, Biology etc.) as well as in engineering sciences (energy, mechanics, electricity, optics etc.). According to projects research objectives, the general object of analysis is represented by the controlled systems with distributed parameters which are usually met in technology dynamical systems with parameter space variation along a single space variable. The standard physical phenomena that are modeled are: diffusion, transport and propagation, thus leading to partial differential equations of parabolic (diffusion), hyperbolic (propagation) and advection first order (transportation), respectively. According to the project main subject, the main application of the studies aimed at the automatically controlled processes in the field of energy in a domain where propagation phenomena are dominant. The scientific novelty of the studies arises from the consideration of the systems described by conservation laws in the following fields: oil drilling and extraction, co-generation (combined heat-electricity generation), energy production in hydroelectric plants.

Program: **PHC CARLOS J FINLEY (Cuba)**

Project title: Modélisation et commande pour les processus de cryoconservation.

Duration: June 2017 - December 2018

Coordinator: Sorin Olaru (France), Marcos Martinez Montero (Turkey).

Abstract: The aim of this project is to initiat a collaboration on subjects related to the mathematical modelling of the dynamics involved in the cryopreservations process. In particular, the viability analysis of the vegetal material subject to cryogeny is one of the main objectives. The approach will realy on the evaluation electric leakage properties.

7.3. International Initiatives

7.3.1. Inria International Partners

7.3.1.1. Informal International Partners

- Louisiana State University, Baton Rouge, USA

- School of Electrical Engineering at the Tel-Aviv University, Israel
- The University of Texas at Austin, Dept. of Aerospace Engineering & Engineering Mechanics, USA
- Blikent University, Turkey
- Universidad de Chile, Chile
- School of Mathematics, University of Leeds, U.K.
- University Federale Rio de Janeiro, Brazil
- UNICAMP, Brazil
- Kyoto University, Japan

7.3.2. Participation in Other International Programs

Giorgio Valmorbida is leading the CNRS funded IRN - SPaDisCo (International Research Network - Systèmes à Paramètres Distribués et Contraintes) composed by more than 50 researchers from Belgium, Czech Republic, France, Italy, Sweden, Turkey and the United Kingdom.

7.4. International Research Visitors

7.4.1. Visits of International Scientists

Stefanella Boatto, Federale University Rio de Janeiro, Brazil, 1 January- 2 March.

Pedro Luis Dias Peres, UNICAMP, Brazil, December 2018.

Valter Leite Junior, CEFET-MG, Brazil, May 2018.

Antonis Papachristodoulou, University of Oxford, UK, September 2018.

Yutaka Yamamoto, Kyoto University, Japan, 6 Sept - 6 Nov.

7.4.2. Visits to International Teams

7.4.2.1. Research Stays Abroad

Islam Boussaada spent one month during July 2018 as a Research Associate at the Department of Electronic Engineering at City University of Hong Kong. He started a research collaboration with Professor Jie Chen Team.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events Organisation

8.1.1.1. Member of the Organizing Committees

Catherine Bonnet is member of the organizing committee of SIAM CT19, Chendu China, July 2019. Islam Boussaada is co-animator of the national working group GT OSYDI of the GDR MACS funded by CNRS. Giorgio Valmorbida was member of the Organizing committee of the 2nd Workshop on DElays and COnstraints on Distributed parameter systems (DECOD-2018).

8.1.2. Scientific Events Selection

- Catherine Bonnet was Associate Editor for the conferences 2019 American Control Conference, Philadelphia, USA.
- Islam Boussaada was Associate Editor for 14th IFAC Workshop on Time Delay Systems, Budapest, Hungary.
- Frédéric Mazenc was Associate Editor for the conferences 2019 American Control Conference, Philadelphie, USA, and the 57th IEEE Conference on Decision and Control, Miami, USA, (2018).

8.1.2.1. Member of the Conference Program Committees

- Catherine Bonnet is a member of the scientific committee of the GDRI (International Research Group funded by CNRS) SpaDisco since 2017.
- Catherine Bonnet, Islam Boussaada and Sorin Olaru are members of the International Program Committee of the Joint IFAC Conference 7th Symposium on System Structure and Control (SSSC 2019) and 15th IFAC Workshop on Delay Systems, Sinaia, Roumania, Sept 2019.
- Islam Boussaada was a member of the International Program Committee for 14th IFAC Workshop on Time Delay Systems (TDS 2018), Budapest, Hungary.
- Frédéric Mazenc and Giorgio Valmorbida are members (Associate Editors) of the *Control Editorial Board - IEEE - CSS*. Sorin Olaru was member of the International Program Committee of the 9th IFAC Symposium on Robust Control Design - ROCOND and of the IEEE Mediterranean Control Conference.

8.1.2.2. Reviewer

The team reviewed papers for several international conferences including IEEE Conference on Decision and Control, IEEE American Control Conference, European Control Conference, IFAC World Congress.

8.1.3. Journal

8.1.3.1. Member of the Editorial Boards

Frédéric Mazenc is member of the editorial boards (Associate Editor) of the following journals:

- IEEE Transactions on Automatic Control;
- European Journal of Control;
- Journal of Control and Decision.

Sorin Olaru is is member of the editorial boards (Associate Editor) of the following journals:

- IEEE Control Systems Letters;
- IMA Journal of Mathematical Control and Information, Oxford Press.

8.1.3.2. Reviewer - Reviewing Activities

The team reviewed papers for several journals including SIAM Journal on Control and Optimization, Automatica, IEEE Transactions on Automatic Control, IEEE Control Systems Magazine, Systems and Control Letters.

8.1.4. Invited Talks

Stefanella Boatto gave a talk at the biomathematics Seminar of Institut de Mathématiques de Marseille, Aix-Marseille univerité, February 2018, Marseilles, France. Title of her talk: '*Modelling epidemics dynamics due to Aedes mosquitoes : the example of Rio de Janeiro. How to approximate an epidemic attractor and to estimate the infectivity rate*'.

Frédéric Mazenc was a plenary speaker of the conference POSTA2018, August 2018, Hangzhou, China. Title of his talk: '*Stability of Positive Systems With Delay: Changes of Coordinates, Comparison Systems, Lyapunov Functionals*'.

Giorgio Valmorbida gave an invited talk at the International Workshop on Robust LPV Control Techniques and Anti-Windup Design, April 2018, Toulouse, France. Title of his talk: '*Anti-Windup Design for Synchronous Machines*'.

Giorgio Valmorbida gave an invited talk at the Meeting of the GT CSE (groupe de travail Commande de Systèmes Éléctriques) , May 2018, Paris, France. Title of his talk: '*Anti-Windup Design for Synchronous Machines*'.

Giorgio Valmorbida gave an invited talk at the 2nd Workshop on Stability and Control of Infinite-Dimensional Systems (SCINDIS-2018), May 2018, Wurzburg, Germany. Title of his talk: '*Convex Optimization Methods to Solve Integral Inequalities*'.

Giorgio Valmorbida gave an invited talk at the 2nd Workshop on DElays and COnstraints on Distributed parameter systems (DECOD-2018), November 2018, Toulouse, France. Title of his talk: '*Stability Analysis of Piece-wise Affine Discrete-Time systems*'.

8.1.5. Leadership within the Scientific Community

Catherine Bonnet is a member of the IFAC Technical Committees on *Distributed Parameter Systems* and on *Biological and Medical Systems*. She is a member of the management committee of the COST Action FRACTAL (2016-2020). Sorin Olaru is a member of the IFAC Technical Committees on *Robust Control* and on *Optimal control*. He is a member of the IEEE Technical Committee on *Hybrid Systems*. He is a Senior Member IEEE since 2012.

8.1.6. Scientific Expertise

Catherine Bonnet is a member of the Evaluation Committee of Inria since September 2015.

Since 2014, Frédéric Mazenc is an expert for the FNRS (Belgium). His mission consists in evaluating research projects funded by this institution.

Since 2012, Frédéric Mazenc is an expert for the ANVUR (National Agency for the Evaluation of Universities and Research Institutes, Italy). His mission consists in evaluating the contribution of Italian scientists.

Since 2011, Frédéric Mazenc is an expert for the Romanian National Council for Development and Innovation (Romania). His mission consists in evaluating research projects funded by the this institution.

8.1.7. Research Administration

Catherine Bonnet is a member of the administration council of the association *Femmes et Mathématiques*, of the *Parity Committee* of Inria and of the *Cellule veille et prospective* of Inria (both created in 2015). She is a member of the Bureau du Comité des Projets du CRI Saclay-Ile-de-France.

In 2018, Frédéric Mazenc is a member of the commission scientifique du CRI Saclay-Ile-de-France. Since October 2017, he is Correspondant Inria Saclay A.M.I.E.S., http://www.agence-maths-entreprises.fr/

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Master : Stefanella Boatto, Challenges in Biomathematical Modelling, 3h, M1, CentraleSupélec, France.

Master : Catherine Bonnet, Stability of Delay Systems, 1.5h, M1, CentraleSupélec.

Licence : Islam Boussaada, Complex analysis, 60, L3, IPSA, France.

Licence : Islam Boussaada, Harmonic analysis, 60, L3, IPSA.

Licence : Sorin Olaru, Automatic Control, 8h , M1, CentraleSupélec.

Licence : Sorin Olaru, Signals and systems, 8h , L3, CentraleSupélec.

Licence : Sorin Olaru, Embedded systems, 8h , M1, CentraleSupélec.

Licence : Sorin Olaru, Numerical methods and Optimization, 24h, M1, CentraleSupélec.

Licence : Sorin Olaru, Hybrid systems, 16h, M2, CentraleSupélec.

Licence : Guillaume Sandou, Signals and Systems, 87h, L3, CentraleSupélec.

Licence : Guillaume Sandou, Model representation and analysis, 70h, L3, CentraleSupélec.

Licence : Guillaume Sandou, Mathematics and programming, 18h, L3, CentraleSupélec.

Licence : Giorgio Valmorbida, Embedded systems, 7.5h, M1, Cursus Ingénieur CentraleSupélec.

Master : Guillaume Sandou, Automatic Control, 8h, M1, CentraleSupélec.

Master : Guillaume Sandou, Numerical methods and optimization, 28h, M1 and M2, CentraleSupélec.

Master : Guillaume Sandou, Modelling and system stability analysis, 21h, M2, CentraleSupélec.

Master : Guillaume Sandou, Control of energy systems, 22h, M2, CentraleSupélec.

Master : Guillaume Sandou, Robust control and mu-analysis, 9h, M2, CentraleSupélec.

Master : Guillaume Sandou, Systems identification, 32h, M2, ENSTA.

Master : Guillaume Sandou, System Analysis, 22h, M2, Ecole des Mines de Nantes.

Master : Giorgio Valmorbida, Numerical methods and optimisation, 6 h, 6hETD, niveau M1, Cursus Ingénieur CentraleSupélec.

Master : Giorgio Valmorbida, Commande d'Entraînements de Vitesse Variable 12 h, M1, Cursus Ingénieur CentraleSupélec.

Master : Giorgio Valmorbida, Control Theory, 7.5 h, M1, Cursus Ingénieur CentraleSupélec.

Master : Giorgio Valmorbida, Dynamical Systems, 9 h, M2, Master Automatique Taitement du Signal et de I'image - Université Paris-Saclay, France.

Master : Giorgio Valmorbida, Control 30h, niveau M1, Master Nuclear Energy - Université Paris-Saclay.

Doctorat : Islam Boussaada, Introduction to the qualitative theory of functional differential equations, 12h, University Mouloud Mammeri, Algeria

8.2.2. Supervision

PhD : Saeed Ahmed, title: *Observer Design and Output Feedback Stabilization of Time Varying Systems*, Bilkent University. Thesis defense: 03 July 2018, supervisors: Hitay Ozbay, Frédéric Mazenc.

PhD in progress : Nadine Aoun, Modélisation de réseaux de chaleur et gestion avancée multi-échelles de la production, de la distribution et de la demande. Modeling and multi-scale advanced management of production, distribution and demand in district heating networks. Supervisor: Guillaume Sandou.

PhD in progress : Leonardo Broering Groff, Periodic Event-Triggered Control, mars 2016. Supervisors: Giorgio Valmorbida and Joao Manoel Gomes da Silva Jr.

PhD in progress : Caetano Cardeliquio, Stability and stabilization of (possibly fractional) systems with delays. French Supervisor : Catherine Bonnet, Brazilian Supervisor : André Fioravanti.

PhD in progress : Mohamed Lotfi Derouiche, Sur l'optimisation par métaheuristiques avancées de lois de commande prédictive non linéaire. On the optimization of nonlinear predictive control laws using advanced metaheuristics algorithms. Supervisors: Soufienne Bouallegue, Joseph Haggége et Guillaume Sandou.

PhD : Nicolo Gionfra, Optimisation du pilotage d'un parc d'énergies renouvelables avec stockage et du réseau de distribution sous-jacent. Optimization of the control of a park of renewable sources considering strorage means and distribution network. Supervisors: Houria Siguerdidjane et Guillaume Sandou. Defended in March 2018.

PhD in progress : Jean Mercat, Modele predictif des objets d'une scene routiere ; application à la sélection robuste des cibles pour les ADAS. Supervisor: Guillaume Sandou.

PhD in progress : Maxime Pouilly-Cathelain, Commande adaptative temps réel vis-a-vis de critères multiples de haut niveau. Supervisor : Guillaume Sandou.

8.2.3. Juries

- Catherine Bonnet was President of the Saclay-Ile-de-France Junior Researcher recruiting committee and a member of the National Inria Junior Researcher recruiting committee.
- Catherine Bonnet was reviewer of the PhD thesis of Mohamad Taki Asghar *'New advanced control strategies for steel making process'*, 3 May 2018, Université de Nancy. She was President of the PhD Defense juries of Noussaiba Gasmi *'Observation et commande des systèmes dynamiques d'ordre non entier'*, 14 November 2018, Université de Nancy, of Pierre-Marie Gibert *'Use of sinusoidal predictors for time-domain simulation of AC power systems'*, 30 November 2018, Université de Lyon, of Abdelkrim Bahloul *'Sur la commande des robots manipulateurs industriels en co-manipulation robotique'*, 7 December 2018, L2S, CentraleSupelec, and member of the PhD Defense jury of Bainan Liu *'Boundary obeserver-based output feedback control of coupled parabolic PDEs'*, 17 December 2018, INSA centre Val de Loire.
- Frédéric Mazenc was a reviewer of the Phd thesis of Ricardo Sanz Diaz, '*Robust control strategies for unstable systems with input/output delays*', September 27, 2018, Universitat Politecnica de Valencia, Spain and of the Phd thesis of Mohammed Safi, '*Stabilité de Lyapunov de systèmes couplés impliquant une équation de transport*', October 31, 2018, LAAS, Université de Toulouse, France.
- Sorin Olaru was President of the PhD Defense jury of Dominique Monnet, '*Global Minmax optimization for robust Hinf control*', université de Brest, France, reviewer of the PhD Defense jury of Nadia Paola Rosero Ibarra,*'Modeling and Observation applied to physiology-aware control for cycling'*, 12 November 2018, université Grenoble Alpes, France and member of the PhD Defense jury of Nassim Loukkas, *'State-membership state observer design based on explicit characterizations of the estimation-error bounds'*, 6 June 2018, université Grenoble Alpes.
- Giorgio Valmorbida was a member of the jury of the Phd thesis of Fabien Niel, *'Modeling and control of a wing at low Reynolds number with high amplitude aeroelastic oscillations'*, 26 January 2018, LAAS, université de Toulouse, France.

8.3. Popularization

8.3.1. Interventions

Catherine Bonnet awarded the price of Mathematics and Italian et Italien at the *Concours Général des lycées et des métiers du Ministère de l'Éducation nationale*, Grand Amphitéatre de La Sorbonne, July 2018. She welcomed a group of middle school students, 17 December 2018.

9. Bibliography

Major publications by the team in recent years

[1] C. BONNET, A. R. FIORAVANTI, J. R. PARTINGTON.*Stability of Neutral Systems with Commensurate Delays and Poles Asymptotic to the Imaginary Axis*, in "SIAM Journal on Control and Optimization", March 2011, vol. 49, nº 2, p. 498-516, https://hal.inria.fr/hal-00782325

[2] C. BONNET, J. PARTINGTON.*Stabilization of some fractional delay systems of neutral type*, in "Automatica", 2007, vol. 43, p. 2047–2053

[3] M. MALISOFF, F. MAZENC.*Constructions of Strict Lyapunov Functions*, Communications and Control Engineering Series, Springer-Verlag London Ltd., 2009

[4] F. MAZENC, M. MALISOFF, S.-I. NICULESCU.*Reduction Model Approach for Linear Time-Varying Systems with Delays*, in "IEEE Transactions on Automatic Control", 2014, vol. 59, n^o 8, p. 2068–2014

[5] W. MICHIELS, S.-I. NICULESCU.*Stability and Stabilization of Time-Delay Systems. An Eigenvalue-Based Approach*, Advances in Design and Control, SIAM: Philadelphia, 2007, vol. 12

[6] S.-I. NICULESCU.*Delay Effects on Stability: a Robust Control Approach*, Lecture Notes in Control and Information Sciences, Springer, 2001, vol. 269

[7] S. OLARU, D. DUMUR.*Avoiding constraints redundancy in predictive control optimization routines*, in "IEEE Trans. Automat. Control", 2005, vol. 50, n^o 9, p. 1459–1465

[8] G. SANDOU.*Particle swarm optimization: an efficient tool for the design of automatic control law*, in "European Control Conference", Budapest, Hungary, August 23rd-26th 2009

Publications of the year

Articles in International Peer-Reviewed Journal

[9] P. AHMADI-MOSHKENANI, T. A. JOHANSEN, S. OLARU.*Combinatorial Approach towards Multi-Parametric Quadratic Programming based on Characterizing Adjacent Critical Regions*, in "IEEE Transactions on Automatic Control", 2018, 1 [*DOI :* 10.1109/TAC.2018.2791479], https://hal-centralesupelec.archives-ouvertes.fr/hal-01720260

[10] S. AHMED, F. MAZENC, H. OZBAY.*Dynamic Output Feedback Stabilization of Switched Linear Systems with Delay via a Trajectory Based Approach* ⋆, in "Automatica", July 2018, vol. 93, p. 92-97 [*DOI :* 10.1016/J.AUTOMATICA.2018.03.072], https://hal.inria.fr/hal-01849039

[11] S. AMRANE, F. BEDOUHENE, I. BOUSSAADA, S.-I. NICULESCU.*On Qualitative Properties of Low-Degree Quasipolynomials: Further remarks on the spectral abscissa and rightmost-roots assignment*, in "Bull. Math. Soc. Sci. Math. Roumanie Tome", 2018, vol. 61, p. 361 - 381, https://hal-centralesupelec.archives-ouvertes.fr/hal-01969066

[12] I. BOUSSAADA, S.-I. NICULESCU.*On the Dominancy of Multiple Spectral Values for Time-delay Systems with Applications*, in "IFAC-PapersOnLine", 2018, vol. 51, n^o 14, p. 55-60, https://hal-centralesupelec.archives-ouvertes.fr/hal-01957523

[13] I. BOUSSAADA, S. TLIBA, S.-I. NICULESCU, H. U. UNAL, T. VYHLÍDAL.*Further remarks on the effect of multiple spectral values on the dynamics of time-delay systems. Application to the control of a mechanical system*, in "Linear Algebra and its Applications", 2018, p. 1-16 [*DOI :* 10.1016/J.LAA.2017.11.022], https://hal-centralesupelec.archives-ouvertes.fr/hal-01657659

[14] D. BRESCH-PIETRI, F. MAZENC, N. PETIT.*Robust compensation of a chattering time-varying input delay with jumps*, in "Automatica", 2018, vol. 92, p. 225-234 [*DOI :* 10.1016/J.AUTOMATICA.2018.03.058], https://hal.inria.fr/hal-01849024

[15] W. Djema, C. Bonnet, F. Mazenc, J. Clairambault, E. Fridman, P. Hirsch, F. Delhommeau.*Control in dormancy or eradication of cancer stem cells: Mathematical modeling and stability issues*, in "Journal of Theoretical Biology", July 2018, vol. 449, p. 103 - 123 [*DOI :* 10.1016/j.jtbi.2018.03.038], https://hal.inria.fr/hal-01852154

[16] H. Dong, Q. Hu, M. Akella, F. Mazenc.*Partial Lyapunov Strictification: Dual Quaternion based Observer for 6-DOF Tracking Control*, in "IEEE Transactions on Control Systems Technology", September 2018, vol. 14, p. 1-17 [*DOI :* 10.1109/tcst.2018.2864723], https://hal.inria.fr/hal-01918284

[17] R. Drummond, G. Valmorbida, S. Duncan.*Generalized Absolute Stability Using Lyapunov Functions With Relaxed Positivity Conditions*, in "IEEE Control Systems Letters", April 2018, vol. 2, n° 2, p. 207 - 212 [*DOI :* 10.1109/LCSYS.2017.2782747], https://hal.archives-ouvertes.fr/hal-01710273

[18] A. Golovkina, D. Ovsyannikov, S. Olaru.*Performance optimization of radioactive waste transmutation in accelerator driven system*, in "Cybernetics and Physics", 2018, vol. 7, n° 4, p. 210 - 215, https://hal-centralesupelec.archives-ouvertes.fr/hal-01968893

[19] M. Gulan, G. Takacs, N. A. Nguyen, S. Olaru, P. Rodriguez-Ayerbe, B. Rohal'-Ilkiv.*Efficient Embedded Model Predictive Vibration Control via Convex Lifting*, in "IEEE Transactions on Control Systems Technology", 2018, vol. PP, n° 99, p. 1-15 [*DOI :* 10.1109/TCST.2017.2764019], https://hal-centralesupelec.archives-ouvertes.fr/hal-01720261

[20] M. Hovd, S. Olaru.*Parameter-dependent PWQ Lyapunov function stability criteria for uncertain piecewise linear systems*, in "Modeling, Identification and Control", 2018, vol. 39, n° 1, p. 15 - 21 [*DOI :* 10.4173/mic.2018.1.2], https://hal-centralesupelec.archives-ouvertes.fr/hal-01736928

[21] D. Irofti, K. Gu, I. Boussaada, S.-I. Niculescu.*Some insights into the migration of double imaginary roots under small deviation of two parameters*, in "Automatica", 2018, vol. 88, p. 91-97 [*DOI :* 10.1016/j.automatica.2017.11.015], https://hal.archives-ouvertes.fr/hal-01712829

[22] C. Jin, K. Gu, I. Boussaada, S.-I. Niculescu.*Stability Analysis of a More General Class of Systems with Delay-Dependent Coefficients*, in "IEEE Transactions on Automatic Control", January 2019, vol. 14, 1, https://hal.archives-ouvertes.fr/hal-01957472

[23] M. Kuře, T. Vyhlídal, W. Michiels, I. Boussaada.*Spectral design of robust delayed resonator by double-root assignment*, in "IFAC-PapersOnLine", 2018, vol. 51, n° 14, p. 72-77, https://hal-centralesupelec.archives-ouvertes.fr/hal-01957517

[24] T. Lathuilière, G. Valmorbida, E. Panteley.*Limit cycles in Liénard systems with saturation*, in "IFAC-PapersOnLine", 2018, vol. 51, n° 33, p. 127-131, https://hal.archives-ouvertes.fr/hal-01974946

[25] T. Lathuilière, G. Valmorbida, E. Panteley.*Periodic orbits in planar linear systems with input saturation*, in "IEEE Control Systems Letters", May 2018, vol. 2, n° 3, p. 435-440 [*DOI :* 10.1109/LCSYS.2018.2841807], https://hal.archives-ouvertes.fr/hal-01811047

[26] A. Martínez-González, C.-F. Méndez-Barrios, S.-I. Niculescu, J. Chen, L. Félix.*Weierstrass approach to asymptotic behavior characterization of critical imaginary roots for retarded differential equations*, in "SIAM Journal on Control and Optimization", January 2019, vol. 57, n° 1, p. 1-22, https://hal.archives-ouvertes.fr/hal-01973973

[27] F. MAZENC, S. AHMED, M. MALISOFF.*Finite Time Estimation Through a Continuous-Discrete Observer*, in "International Journal of Robust and Nonlinear Control", October 2018, vol. 28, n^{o} 16, p. 4831-4849 [*DOI :* 10.1002/RNC.4286], https://hal.inria.fr/hal-01918275

[28] F. MAZENC, M. MALISOFF, L. BURLION, J. WESTON.*Bounded Backstepping Control and Robustness Analysis for Time-Varying Systems under Converging- Input-Converging-State Conditions* *, in "European Journal of Control", July 2018, vol. 42, p. 15-24 [*DOI :* 10.1016/J.EJCON.2018.02.005], https://hal.inria.fr/hal-01848927

[29] F. MAZENC, M. MALISOFF, H. OZBAY.*Stability and robustness analysis for switched systems with time-varying delays*, in "SIAM Journal on Control and Optimization", January 2018, vol. 56, n^{o} 1, p. 158-182 [*DOI :* 10.1137/16M1104895], https://hal.inria.fr/hal-01660106

[30] F. MAZENC, M. MALISOFF, H. ¨. OZBAY.*Stability Analysis of Switched Systems with Time-Varying Discontinuous Delays* *, in "SIAM Journal on Control and Optimization", 2018, https://hal.inria.fr/hal-01849046

[31] F. MAZENC, G. ROBLEDO, M. MALISOFF.*Stability and Robustness Analysis for a Multispecies Chemostat Model with Delays in the Growth Rates and Uncertainties*, in "Discrete and Continuous Dynamical Systems - Series B", April 2018, https://hal.inria.fr/hal-01918273

[32] H. MOUNIER, S.-I. NICULESCU, A. CELA, M. STEFAN GEAMANU.*Flatness-based longitudinal vehicle control with embedded torque constraint*, in "IMA Journal of Mathematical Control and Information", February 2018, p. 1-16 [*DOI :* 10.1093/IMAMCI/DNY005], https://hal.archives-ouvertes.fr/hal-01705876

[33] S. MUNIR, M. HOVD, S. OLARU.*Low complexity constrained control using higher degree Lyapunov functions*, in "Automatica", December 2018, vol. 98, p. 215 - 222 [*DOI :* 10.1016/J.AUTOMATICA.2018.09.030], https://hal-centralesupelec.archives-ouvertes.fr/hal-01908231

[34] N. A. NGUYEN, M. GULAN, S. OLARU, P. RODRIGUEZ-AYERBE.*Convex Liftings: Theory and Control Applications*, in "IEEE Transactions on Automatic Control", 2018, 1 [*DOI :* 10.1109/TAC.2017.2737234], https://hal-centralesupelec.archives-ouvertes.fr/hal-01720262

[35] L. H. V. NGUYEN.*A unified approach for the H_∞-stability analysis of classical and fractional neutral systems with commensurate delays*, in "SIAM Journal on Control and Optimization", 2018, vol. 56, n^{o} 1, p. 538-555 [*DOI :* 10.1137/16M1101271], https://hal.archives-ouvertes.fr/hal-01679070

[36] S. I. NICULESCU, C. JIN, K. GU, I. BOUSSAADA.*Stability Analysis of Systems with Delay-Dependent Coefficients: an Overview*, in "IEEE Access", April 2018, 1 [*DOI :* 10.1109/ACCESS.2018.2828871], https://hal-centralesupelec.archives-ouvertes.fr/hal-01781318

[37] S. OLARU, H. ITO.*Characterization of ultimate bounds for systems with state-dependent disturbances*, in "IEEE Control Systems Letters", 2018, vol. 2, n^{o} 4, p. 797-802 [*DOI :* 10.1109/LCSYS.2018.2849642], https://hal-centralesupelec.archives-ouvertes.fr/hal-01830572

[38] X.-M. SUN, X.-F. WANG, F. MAZENC.*Delay-hybrid-dependent stability for systems with large delays*, in "IEEE transactions on systems, man, and cybernetics", 2018, p. 1-12, https://hal.inria.fr/hal-01849022

[39] G. VALMORBIDA, R. DRUMMOND, S. DUNCAN.*Regional Analysis of Slope Restricted Lur'e Systems*, in "IEEE Transactions on Automatic Control", June 2018, 1, https://hal.archives-ouvertes.fr/hal-01969685

[40] Y. WANG, R. RAJAMANI, A. ZEMOUCHE.*Sequential LMI approach for design of a BMI-based robust observer state feedback controller with nonlinear uncertainties*, in "International Journal of Robust and Nonlinear Control", March 2018, vol. 28, n^{o} 4, p. 1246-1260 [*DOI :* 10.1002/RNC.3948], https://hal.archives-ouvertes.fr/hal-01575709

Invited Conferences

[41] S. BOATTO, C. BONNET, B. CAZELLES, F. MAZENC, L. H. V. NGUYEN.*Modelling epidemics dynamics due to Aedes mosquitoes : the example of Rio de Janeiro. How to approximate an epidemic attractor and to estimate the infectivity rate*, in "MB2 3rd Mathematical Biology Modeling days of Besançon", Besançon, France, June 2018, https://hal.inria.fr/hal-01977053

International Conferences with Proceedings

[42] J. ALVAREZ-MUNOZ, J. ESCAREÑO, F. MENDEZ-BARRIOS, I. BOUSSAADA, I. NICULESCU, D. NIETO-HERNRANDEZ.*Time-delay Tolerant Control of an Omnidirectional Multi-agent System for Transport Operations*, in "ICSTCC", SINAIA, Romania, October 2018, https://hal.archives-ouvertes.fr/hal-01969082

[43] C. BENNANI, F. BEDOUHENE, A. ZEMOUCHE, H. BIBI, K. CHAIB DRAA, A. AITOUCHE, R. RAJAMANI.*Robust H_∞ observer-based stabilization of linear discrete-time systems with parameter uncertainties*, in "American Control Conference, ACC 2018", Milwaukee, WI, United States, June 2018 [*DOI :* 10.23919/ACC.2018.8431745], https://hal.archives-ouvertes.fr/hal-01736812

[44] F. BLANCHINI, D. CASAGRANDE, G. GIORDANO, S. MIANI, S. OLARU, V. REPPA.*Fault Isolation for Large Scale Discrete-Time Systems Based on Implicit Set Representation*, in "European Control Conference (ECC 2018)", Limassol, Cyprus, June 2018, p. 1-6, https://hal-centralesupelec.archives-ouvertes.fr/hal-01720264

[45] I. BOUSSAADA, S.-I. NICULESCU, K. TRABELSI.*Towards a Decay Rate Assignment Based Design for Time-Delay Systems with Multiple Spectral Values*, in "MTNS", Hong Kong, China, July 2018, https://hal-centralesupelec.archives-ouvertes.fr/hal-01957555

[46] C. CARDELIQUIO, A. FIORAVANTI, C. BONNET, S.-I. NICULESCU.*Stability and Robust Stabilisation Through Envelopes for Retarded Time-Delay Systems*, in "9th IFAC Symposium on Robust Control Design", Florianopolis, Brazil, September 2018, https://hal.inria.fr/hal-01970655

[47] D. DANCIU, I. BOUSSAADA, F. STNG.*Computational modeling and oscillations damping of axial vibrations in a drilling system*, in "ICSTCC", SINAIA, Romania, October 2018, https://hal.archives-ouvertes.fr/hal-01969076

[48] W. DJEMA, C. BONNET, F. MAZENC, J. CLAIRAMBAULT.*Introducing Cell-Plasticity Mechanisms into a Class of Cell Population Dynamical Systems*, in "IEEE American Control Conference (ACC 2018)", Milwaukee, United States, June 2018 [*DOI :* 10.23919/ACC.2018.8430758], https://hal.inria.fr/hal-01848890

[49] W. DJEMA, F. MAZENC, C. BONNET, J. CLAIRAMBAULT, E. FRIDMAN.*Stability Analysis of a Nonlinear System with Infinite Distributed Delays Describing Cell Dynamics*, in "IEEE American Control Conference (ACC 2018)", Milwaukee, United States, June 2018 [*DOI :* 10.23919/ACC.2018.8430869], https://hal.inria.fr/hal-01849010

[50] A. Iovine, F. Mazenc.*Bounded Control for DC/DC Converters: Application to Renewable Sources*, in "57th IEEE Conference on Decision and Control", Miami, United States, December 2018, https://hal.inria.fr/hal-01918305

[51] M. Malisoff, F. Mazenc.*Sequential Predictors for Linear Time-Varying Systems with Delays in the Vector Field and in the Input*, in "57th IEEE Conference on Decision and Control", Miami, United States, December 2018, https://hal.inria.fr/hal-01918343

[52] A. Martínez-González, C. F. Méndez-Barrios, S. I. Niculescu, J. Chen.*Some Remarks on the Asymptotic Behavior for Quasipolynomials with Two Delays*, in "TDS 2018 - 14th IFAC Workshop on Time Delay Systems", Budapest, Hungary, June 2019, vol. 51, n^{o} 14, p. 318-323 [*DOI :* 10.1016/j.ifacol.2018.07.243], https://hal.inria.fr/hal-01950608

[53] F. Mazenc, S. Ahmed.*Finite Time Estimation via Piecewise Constant Measurements* $\star$, in "Second IFAC Conference on Modelling, identification and Control of Nonlinear Systems", Guadalajara, Mexico, June 2018, https://hal.inria.fr/hal-01848996

[54] F. Mazenc, S. Ahmed, M. Malisoff.*Reduced Order Finite Time Observers for Time-Varying Nonlinear Systems*, in "57th IEEE Conference on Decision and Control", Miami, United States, December 2018, https://hal.inria.fr/hal-01918307

[55] F. Mazenc, L. Burlion, V. Gibert.*Stabilization of a nonlinear system that arises in the context of vision based landing of an airliner*, in "57th IEEE Conference on Decision and Control", Miami, United States, December 2018, https://hal.inria.fr/hal-01918348

[56] F. Mazenc, L. Burlion, V. Gibert.*Stabilization with imprecise measurements: application to a vision based landing problem*, in "IEEE American Control Conference (ACC 2018)", Milwaukee, United States, June 2018 [*DOI :* 10.23919/acc.2018.8431402], https://hal.inria.fr/hal-01848888

[57] F. Mazenc, L. Burlion, M. Malisoff.*Backstepping Design for Output Feedback Stabilization for Uncertain Systems using Dynamic Extension*, in "Second IFAC Conference on Modelling, identification and Control of Nonlinear Systems", Guadalajara, Mexico, June 2018, https://hal.inria.fr/hal-01918299

[58] F. Mazenc, M. Malisoff.*Continuous-Discrete Sequential Observers under Sampling and Input Delays*, in "IEEE American Control Conference (ACC 2018)", Milwaukee, United States, June 2018 [*DOI :* 10.23919/acc.2018.8431518], https://hal.inria.fr/hal-01848887

[59] N. Michel, S. Olaru, G. Valmorbida, S. Bertrand, D. Dumur.*Invariant Sets for Discrete-Time Constrained Linear Systems Using Sliding Mode Approach*, in "European Control Conference (ECC 2018)", Limassol, Cyprus, June 2018, p. 2929-2934 [*DOI :* 10.23919/ECC.2018.8550119], https://hal.archives-ouvertes.fr/hal-01718460

[60] N. A. Nguyen, S. Olaru.*A piecewise affine control Lyapunov function for robust control*, in "European Control Conference (ECC 2018)", Limassol, Cyprus, June 2018, https://hal-centralesupelec.archives-ouvertes.fr/hal-01720263

[61] G. Phanomchoeng, A. Zemouche, W. Jeon, R. Rajamani, F. Mazenc.*An H_∞ observer for descriptor nonlinear systems with nonlinear output equations*, in "American Control Conference, ACC 2018", Milwaukee, WI, United States, June 2018, https://hal.inria.fr/hal-01849018

Other Publications

[62] S. BOATTO, C. BONNET, B. CAZELLES, F. MAZENC.*SIR model with time dependent infectivity parameter : approximating the epidemic attractor and the importance of the initial phase*, January 2018, working paper or preprint, https://hal.inria.fr/hal-01677886

[63] S. BOATTO, C. BONNET, B. CAZELLES, F. MAZENC, L. H. V. NGUYEN.*Modelling epidemics dynamics due to Aedes mosquitoes : the example of Rio de Janeiro. How to approximate an epidemic attractor and to estimate the infectivity rate*, November 2018, DinamicAs, Poster, https://hal.inria.fr/hal-01976151

[64] S. SLIMANI, P. RAYNAUD DE FITTE, I. BOUSSAADA.*Dynamics of a prey-predator system with modified Leslie-Gower and Holling type II schemes incorporating a prey refuge*, January 2019, https://arxiv.org/abs/1802.05967 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01707650

References in notes

[65] H. KITANO.*Biological robustness*, in "Nature", 2004, vol. 5, p. 826–837

[66] J. E. MARSHALL, H. GÓRECKI, A. KORYTOWSKI, K. WALTON.*Time-Delay Systems: stability and performance criteria with applications*, Ellis Horwood, 1992

[67] H. OZBAY, C. BONNET, A. FIORAVANTI.*PID Controller Design for Fractional-Order Systems with Time Delays*, in "Systems and Control Letters", January 2012, vol. 61, n^{o} 1, p. 18-23, http://hal.inria.fr/hal-00766062

Project-Team EX-SITU

Extreme Situated Interaction

IN COLLABORATION WITH: Laboratoire de recherche en informatique (LRI)

IN PARTNERSHIP WITH:
Université Paris-Sud (Paris 11)

RESEARCH CENTER
Saclay - Île-de-France

THEME
Interaction and visualization

Table of contents

Project-Team EX-SITU

Creation of the Team: 2015 January 01, updated into Project-Team: 2017 July 01

Keywords:

Computer Science and Digital Science:

A5.1. - Human-Computer Interaction
A5.1.1. - Engineering of interactive systems
A5.1.2. - Evaluation of interactive systems
A5.1.5. - Body-based interfaces
A5.1.6. - Tangible interfaces
A5.1.7. - Multimodal interfaces

Other Research Topics and Application Domains:

B2.8. - Sports, performance, motor skills
B5.7. - 3D printing
B6.3.1. - Web
B6.3.4. - Social Networks
B9.2. - Art
B9.2.1. - Music, sound
B9.2.4. - Theater
B9.5. - Sciences

1. Team, Visitors, External Collaborators

Research Scientists

Wendy Mackay [Team leader, Inria, Senior Researcher, HDR]
Theophanis Tsandilas [Inria, Researcher]
Baptiste Caramiaux [CNRS, Researcher]

Faculty Members

Joanna Mcgrenere [University of British Columbia and Inria Chair, from May 2018 until Jul 2018]
Michel Beaudouin-Lafon [Univ Paris-Sud, Professor, HDR]
Sarah Fdili Alaoui [Univ Paris-Sud, Assistant Professor]
Cédric Fleury [Univ Paris-Sud, Assistant Professor]

External Collaborator

Midas Nouwens [Aarhus University]

Technical Staff

Romain Di Vozzo [Inria, until May 2018]
Lawrence Fyfe [Inria]
Romain Primet [Inria]
Nicolas Taffin [Inria]

PhD Students

Dimitrios Christaras Papageorgiou [Univ Paris-Sud, from Oct 2018]
Marianela Ciolfi Felice [Univ Paris-Sud]
Carla Griggio [Inria]
Per Carl Viktor Gustafsson [Univ Paris-Sud, from Oct 2018]

Han Han [Univ Paris-Sud, from Oct 2018]
Shu-Yuan Hsueh [Univ Paris-Sud]
Germán Leiva [Inria]
Jiali Liu [Institut Telecom, until Apr 2018]
Wanyu Liu [Institut Telecom]
Yujiro Okuya [Univ Paris-Sud]
Miguel Renom [Univ Paris-Sud, from Oct 2018]
Jean-Philippe Rivière [Univ Paris-Sud]
Téo Sanchez [Inria, from Sep 2018]
Philip Tchernavskij [Univ Paris-Saclay]
Elizabeth Walton [Univ Paris-Sud, from Nov 2018]
Michael Wessely [Inria]
Yi Zhang [Inria, from Sep 2018]
Yiran Zhang [Univ Paris-Sud]

Post-Doctoral Fellows

Jessalyn Alvina [Inria, until May 2018]
Benjamin Bressolette [Univ Paris-Sud, from Dec 2018]
John Maccallum [Inria, until Nov 2018]
Nolwenn Maudet [Inria, until Mar 2018]
Andrew Webb [Inria]

Visiting Scientist

Janin Koch [Aalto University, from Oct 2018]

Administrative Assistant

Alexandra Merlin [Inria]

2. Overall Objectives

2.1. Overall Objectives

Interactive devices are everywhere: we wear them on our wrists and belts; we consult them from purses and pockets; we read them on the sofa and on the metro; we rely on them to control cars and appliances; and soon we will interact with them on living room walls and billboards in the city. Over the past 30 years, we have witnessed tremendous advances in both hardware and networking technology, which have revolutionized all aspects of our lives, not only business and industry, but also health, education and entertainment. Yet the ways in which we interact with these technologies remains mired in the 1980s. The graphical user interface (GUI), revolutionary at the time, has been pushed far past its limits. Originally designed to help secretaries perform administrative tasks in a work setting, the GUI is now applied to every kind of device, for every kind of setting. While this may make sense for novice users, it forces expert users to use frustratingly inefficient and idiosyncratic tools that are neither powerful nor incrementally learnable.

ExSitu explores the limits of interaction — how extreme users interact with technology in extreme situations. Rather than beginning with novice users and adding complexity, we begin with expert users who already face extreme interaction requirements. We are particularly interested in creative professionals, artists and designers who rewrite the rules as they create new works, and scientists who seek to understand complex phenomena through creative exploration of large quantities of data. Studying these advanced users today will not only help us to anticipate the routine tasks of tomorrow, but to advance our understanding of interaction itself. We seek to create effective human-computer partnerships, in which expert users control their interaction with technology. Our goal is to advance our understanding of interaction as a phenomenon, with a corresponding paradigm shift in how we design, implement and use interactive systems. We have already made significant progress through our work on instrumental interaction and co-adaptive systems, and we hope to extend these into a foundation for the design of all interactive technology.

3. Research Program

3.1. Research Program

We characterize Extreme Situated Interaction as follows:

Extreme users. We study extreme users who make extreme demands on current technology. We know that human beings take advantage of the laws of physics to find creative new uses for physical objects. However, this level of adaptability is severely limited when manipulating digital objects. Even so, we find that creative professionals—artists, designers and scientists—often adapt interactive technology in novel and unexpected ways and find creative solutions. By studying these users, we hope to not only address the specific problems they face, but also to identify the underlying principles that will help us to reinvent virtual tools. We seek to shift the paradigm of interactive software, to establish the laws of interaction that significantly empower users and allow them to control their digital environment.

Extreme situations. We develop extreme environments that push the limits of today's technology. We take as given that future developments will solve “practical" problems such as cost, reliability and performance and concentrate our efforts on interaction in and with such environments. This has been a successful strategy in the past: Personal computers only became prevalent after the invention of the desktop graphical user interface. Smartphones and tablets only became commercially successful after Apple cracked the problem of a usable touch-based interface for the iPhone and the iPad. Although wearable technologies, such as watches and glasses, are finally beginning to take off, we do not believe that they will create the major disruptions already caused by personal computers, smartphones and tablets. Instead, we believe that future disruptive technologies will include fully interactive paper and large interactive displays.

Our extensive experience with the Digiscope WILD and WILDER platforms places us in a unique position to understand the principles of distributed interaction that extreme environments call for. We expect to integrate, at a fundamental level, the collaborative capabilities that such environments afford. Indeed almost all of our activities in both the digital and the physical world take place within a complex web of human relationships. Current systems only support, at best, passive sharing of information, e.g., through the distribution of independent copies. Our goal is to support active collaboration, in which multiple users are actively engaged in the lifecycle of digital artifacts.

Extreme design. We explore novel approaches to the design of interactive systems, with particular emphasis on extreme users in extreme environments. Our goal is to empower creative professionals, allowing them to act as both designers and developers throughout the design process. Extreme design affects every stage, from requirements definition, to early prototyping and design exploration, to implementation, to adaptation and appropriation by end users. We hope to push the limits of participatory design to actively support creativity at all stages of the design lifecycle. Extreme design does not stop with purely digital artifacts. The advent of digital fabrication tools and FabLabs has significantly lowered the cost of making physical objects interactive. Creative professionals now create hybrid interactive objects that can be tuned to the user's needs. Integrating the design of physical objects into the software design process raises new challenges, with new methods and skills to support this form of extreme prototyping.

Our overall approach is to identify a small number of specific projects, organized around four themes: *Creativity, Augmentation, Collaboration* and *Infrastructure*. Specific projects may address multiple themes, and different members of the group work together to advance these different topics.

4. Application Domains

4.1. Creative industries

We work closely with creative professionals in the arts and in design, including music composers, musicians, and sound engineers; painters and illustrators; dancers and choreographers; theater groups; game designers; graphic and industrial designers; and architects.

4.2. Scientific research

We work with creative professionals in the sciences and engineering, including neuroscientists and doctors; programmers and statisticians; chemists and astrophysicists; and researchers in fluid mechanics.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

- ANR ELEMENT project was accepted.
- CNRS PEPS project was accepted.
- ERC CREATIV was extended for a year.
- Wanyu Liu, Olivier Rioul, Joanna McGrenere, Wendy Mackay, and Michel Beaudouin- Lafon: Honorable Mention award at ACM CHI 2018 for “BIGFile: Bayesian Information Gain for Fast File Retrieval” [22]

6. New Software and Platforms

6.1. Platforms

6.1.1. WildOS

Participant: Michel Beaudouin-Lafon [correspondant].

WildOS is middleware designed to support applications that run in an interactive room, such as our WILD and WILDER rooms, with various interaction resources, including a tiled wall display, a motion tracking system, interactive tabletops, tablets, smartphones and custom-made or 3d printed interactive devices. The conceptual model of WildOS is a *platform*, such as the WILD or WILDER room, that can be described as a set of devices on which one or more applications can be run.

WildOS consists of a server running on a machine that has network access to all the machines involved in the platform, and a set of clients running on the various interaction resources, such as a display cluster or a tablet. Once *WildOS* is running, applications can be started and stopped and devices can be added to or removed from the platform.

WildOS relies on Web technologies, most notably Javascript and node.js, as well as node-webkit and HTML5. This makes it inherently portable (it is currently tested on Mac OS X and Linux). While applications can be developed only with these Web technologies, it is also possible to bridge to existing applications developed in other environments if they provide sufficient access for remote control. Sample applications include a web browser, an image viewer, a window manager, and the BrainTwister application developed in collaboration with neuroanatomists at NeuroSpin.

WildOS is used for several research projects at ExSitu and by other partners of the Digiscope project. It was also deployed on several of Google’s interactive rooms in Mountain View, Dublin and Paris. It is available under on Open Source licence at https://bitbucket.org/mblinsitu/wildos.

- ACM: H.5.2 [User Interfaces]: Graphical user interfaces (GUI)
- Software benefit: helps development of multisurface applications.
- OS/Middleware: Crossplatform
- Required library or software: node.js, node-webkit
- Programming language: Javascript

6.1.2. Unity Cluster

Participants: Cédric Fleury [correspondant], Olivier Gladin [SED-SAC].

Unity Cluster is middleware to distribute any Unity 3D (https://unity3d.com/) application on a cluster of computers that run in interactive rooms, such as our WILD and WILDER rooms, or immersive CAVES (Computer-Augmented Virtual Environments). Users can interact the the application with various interaction resources.

Unity Cluster provides an easy solution for running existing Unity 3D applications on any display that requires a rendering cluster with several computers. *Unity Cluster* is based on a master-slave architecture: The master computer runs the main application and the physical simulation as well as manages the input; the slave computers receive updates from the master and render small parts of the 3D scene. *Unity Cluster* manages data distribution and synchronization among the computers to obtain a consistent image on the entire wall-sized display surface.

Unity Cluster can also deform the displayed images according to the user's position in order to match the viewing frustum defined by the user's head and the four corners of the screens. This respects the motion parallax of the 3D scene, giving users a better sense of depth.

Unity Cluster is composed of a set of C Sharp scripts that manage the network connection, data distribution, and the deformation of the viewing frustum. In order to distribute an existing application on the rendering cluster, all scripts must be embedded into a Unity package that is included in an existing Unity project.

- ACM: C.2.4 [Distributed Systems]: Distributed applications, I.3.7 [3D Graphics and Realism]: Virtual reality
- Software benefit: adapts existing Unity 3D application to a rendering cluster of an interactive room.
- OS/Middleware: Crossplatform
- Required library or software: Unity 3D
- Programming language: C Sharp

6.1.3. WILDER

Participants: Michel Beaudouin-Lafon [correspondant], Cédric Fleury, Olivier Gladin.

WILDER (Figure 1) is our second experimental ultra-high-resolution interactive environment, which follows the WILD platform developed in 2009. It features a wall-sized display with seventy-five 20" LCD screens, i.e. a 5m50 x 1m80 (18' x 6') wall displaying 14 400 x 4 800 = 69 million pixels, powered by a 10-computer cluster and two front-end computers. The platform also features a camera-based motion tracking system that lets users interact with the wall, as well as the surrounding space, with various mobile devices. The display uses a multitouch frame (the largest of its kind in the world) to make the entire wall touch sensitive.

WILDER was inaugurated in June, 2015. It is one of the ten platforms of the Digiscope Equipment of Excellence and, in combination with WILD and the other Digiscope rooms, provides a unique experimental environment for collaborative interaction.

In addition to using WILD and WILDER for our research, we have also developed software architectures and toolkits, such as WildOS and Unity Cluster, that enable developers to run applications on these multi-device, cluster-based systems.

7. New Results

7.1. Fundamentals of Interaction

Participants: Michel Beaudouin-Lafon [correspondant], Wendy Mackay, Cédric Fleury, Theophanis Tsandilas, Dimitrios Christaras Papageorgiou, Han Han, Germán Leiva, Nolwenn Maudet, Yujiro Okuya, Miguel Renom, Philip Tchernavskij, Andrew Webb.

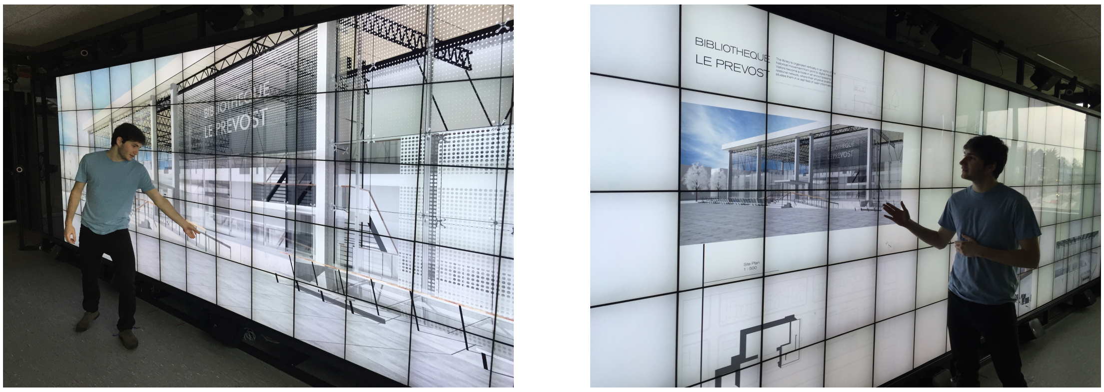


Figure 1. The WILDER platform.

In order to better understand fundamental aspects of interaction, ExSitu conducts in-depth observational studies and controlled experiments which contribute to theories and frameworks that unify our findings and help us generate new, advanced interaction techniques. Our theoretical work also leads us to deepen or re-analyze existing theories and methodologies in order to gain new insights.

Continuing our long-standing exploration of Fitts' law, we demonstrated the dangers of confounding factors in Fitts'-like experimental designs and recommended how to avoid them [20]. Confounds come from the fact that traditional Fitts'-like experiments use geometric progressions of the two main factors (target distance D and amplitude W) and aggregate data points per $ID = log(1 + D/W)$. This typically leads to a strong confound between D and ID, whereby an effect attributed to ID may in fact be due solely to D. We showed evidence of published results where this confound led to the misinterpretation of experimental results, and proposed stochastic sampling of D and W as a technique to avoid such problems.

We also reviewed statistical methods for the analysis of user-elicited gestural vocabularies [16] and argued that current statistics for assessing agreement across participants are problematic. First, we showed that raw agreement rates disregard agreement that occurs by chance and do not reliably capture how participants distinguish among referents. Second, we explained why current recommendations on how to interpret agreement scores rely on incorrect assumptions. Third, we demonstrated that significance tests for comparing agreement rates, either within or between participants, yield large Type I error rates ($> 40\%$ for $\alpha = .05$). As alternatives, we presented agreement indices that are routinely used in inter-rater reliability studies. We discussed how to apply them to gesture elicitation studies. We also demonstrated how to use common resampling techniques to support statistical inference with interval estimates. We applied these methods to reanalyze and reinterpret the findings of four gesture elicitation studies. We also participated in an invited formal debate at ACM/CHI 2018 to discuss the issue of replicability in HCI experiments, specifically whether or not the community should adopt the TOP (Transparency and Openness) guidelines for data and code transparency, citation, experiment preregistration and replication of experiments.

In order to explore novel forms of interaction based on the concepts of *interaction instruments* and *interactive substrates*, we conducted several studies and developed prototypes in three main areas:

First, we challenged the notion of application as the main organizing principle of digital environments. Most of our current interactions with the digital world are mediated by applications that impose artificial limits on collaboration among users and distribution across devices, and the constantly changing procedures that disrupt everyday use. These limitations are due partly to the engineering principles of encapsulation and program-data separation, which highlight the needs for appropriate conceptual models of interaction [18]. We proposed new architectural principles [28], [17] that address these issues by considering interactions as first-class objects that can be dynamically created, added to and removed from an interactive system.

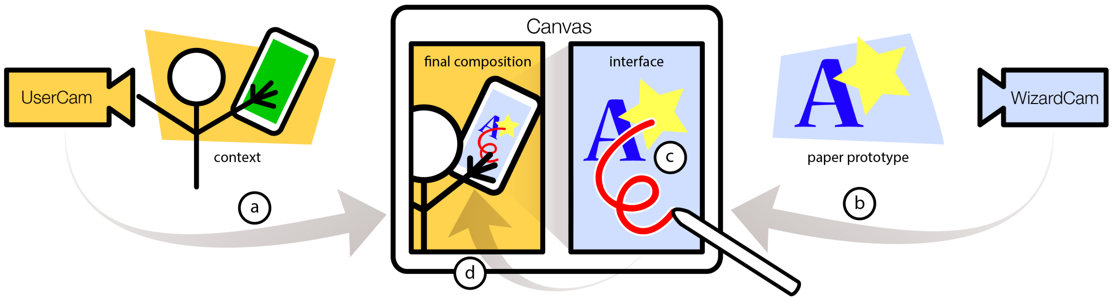


Figure 2. Montage: the UserCam captures the context (a) and the WizardCam captures the paper prototype (b); Both live-stream video to the Canvas, where the designer can add digital sketches (c). Montage replaces the green screen with the interface to create the final composition (d).

Second, we addressed the needs of designers and developers of interactive systems through a series of studies and prototypes. Current prototyping tools do not adequately support the early stages of design, nor the necessary communication between designers and developers. We created and evaluated VideoClipper and Montage [21], two tools that facilitate video prototyping for the early sketching of ideas. VideoClipper facilitates the planning and capturing of video brainstorming ideas and video prototypes, while Montage (fig. 2) uses chroma-keying to create more advanced video prototypes and facilitating their reuse in different contexts. We also created Enact (under submission), a prototyping tool that lets designers and developers work in the same environment to create novel touch-based interaction techniques. Germán Leiva, supervised by Michel Beaudouin-Lafon, successfully defended his Ph.D. thesis *Interactive Prototyping of Interactions: From Throwaway Prototypes to Takeaway Prototyping* [34] on this topic.

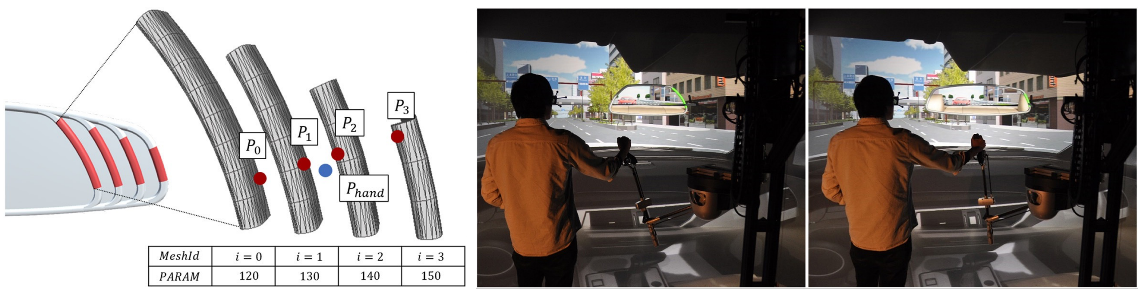


Figure 3. Pre-computed meshes of a rear-view mirror while modifying the right part: the user's hand position (P_{hand}) determines the selected shape (left). A virtual car cockpit where the user modifies the rear-view mirror shape in real time, using haptic force feedback (right).

Third, in the context of Computer Aided Design (CAD), we explored solutions for modifying parametric CAD objects in an immersive virtual reality system. In particular, we developed *ShapeGuide* [14], a technique that lets users modify parameter values by directly pushing or pulling the surface of a CAD object (Figure 3).

Including force feedback increases the precision of the users' hand motions in the 3D space. In a controlled experiment, we compared *ShapeGuide* to a standard one-dimensional scroll technique to measure its added value for parametric CAD data modification on a simple industrial object. We also evaluated the effect of force feedback assistance on both techniques. We demonstrated that *ShapeGuide* is significantly faster and more efficient than the scroll technique. In addition, we showed that force feedback assistance enhances the precision of both techniques.

7.2. Human-Computer Partnerships

Participants: Wendy Mackay [correspondant], Baptiste Caramiaux, Téo Sanchez, Marianela Ciolfi Felice, Carla Griggio, Shu Yuan Hsueh, Wanyu Liu, John Maccallum, Nolwenn Maudet, Joanna Mcgrenere, Midas Nouwens, Andrew Webb.

ExSitu is interested in designing effective human-computer partnerships, in which expert users control their interaction with technology. Rather than treating the human users as the 'input' to a computer algorithm, we explore human-centered machine learning, where the goal is to use machine learning and other techniques to increase human capabilities. Much of human-computer interaction research focuses on measuring and improving productivity: our specific goal is to create what we call 'co-adaptive systems' that are discoverable, appropriable and expressive for the user.

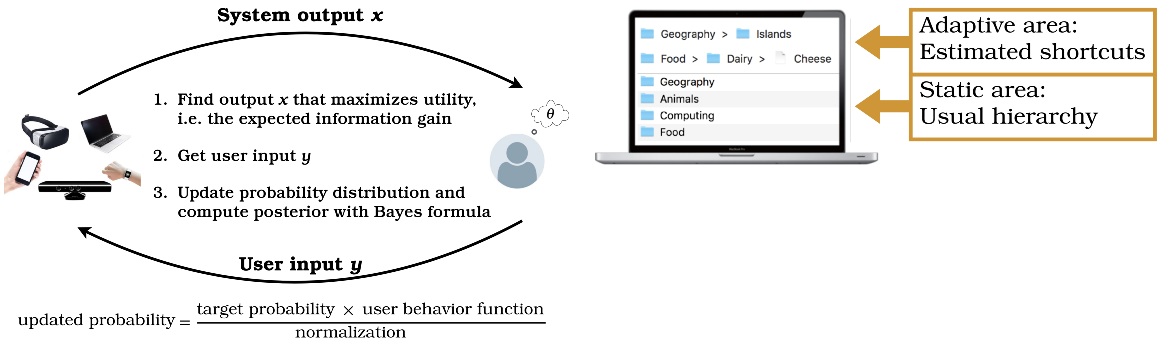


Figure 4. The BIG framework (left) and its application to the BIGFile split-adaptive interface for file navigation (right).

The *Bayesian Information Gain* (BIG) project uses Bayesian Experimental Design, where the criterion is to maximize the information-theoretic concept of mutual information, also known as information gain (fig. 4-left). The resulting interactive system "runs experiments" on the user in order to maximize the information gain from the user's next input and get to the user's goal more efficiently. *BIGnav* applies BIG to multiscale navigation [7]. Rather than simply executing the navigation commands issued by the user, *BIGnav* interprets them to update its knowledge about the user's intended target, and then computes a new view that maximizes the expected information gain provided by the user's next input. This view is located such that, from the system's perspective, the possible navigation commands are uniformly probable, to the extent possible. *BIGFile* [22] (ACM CHI Honorable Mention award) uses a similar approach for file navigation, with a split interface (fig. 4-right) that combines a classical area where users can navigate the file system as usual and an adaptive area with a set of shortcuts calculated with BIG. *BIGnav* and *BIGFile* create a novel form of human-computer partnership, where the computer challenges the user in order to extract more information from the user's input, making interaction more efficient. We showed that both techniques are significantly faster (40% and more) than conventional navigation techniques. Wanyu Liu, supervised by Michel Beaudouin-Lafon, successfully defended her Ph.D. thesis *Information theory as a unified tool for understanding and designing human-computer interaction* [35] on this topic.

In the area of visualization, we studied the common challenge faced by domain experts when identifying and comparing patterns in time series data. While automatic measures exist to compute time series similarity, human intervention is often required to visually inspect these automatically generated results. In collaboration with the ILDA Inria team and Univ. Paris-Descartes, we studied how different visualization techniques affect similarity perception in EEG signals [12], [31]. Our goal was to understand if the time series results returned from automatic similarity measures are perceived in a similar manner, irrespective of the visualization technique; and if what people perceive as similar with each visualization aligns with different automatic measures and their similarity constraints. Overall, our work indicates that the choice of visualization affects which temporal patterns we consider to be similar, i.e., the notion of similarity in a time series is not visualization independent. This demonstrates the need for effective human-computer partnerships in which the computer complements, rather than replaces, human skills and expertise.

We began to explore *human-centred machine learning*, which takes advantage of *active machine learning* to facilitate personalization of an interactive system. We developed a gesture-based recognition system where the user iteratively provides instances and also answers the system's queries. Our results demonstrated the phenomenon of co-adaptation between the human user and the system, which challenges the state of the art in conventional active learning. We further explored interactive reinforcement learning as a way to explore high-dimensional parametric space efficiently [24].

7.3. Creativity

Participants: Sarah Fdili Alaoui [correspondant], Marianela Ciolfi Felice, Carla Griggio, Shu Yuan Hsueh, Germán Leiva, John Maccallum, Wendy Mackay, Baptiste Caramiaux, Nolwenn Maudet, Joanna Mcgrenere, Midas Nouwens, Jean-Philippe Rivière, Nicolas Taffin, Philip Tchernavskij, Theophanis Tsandilas, Andrew Webb, Michael Wessely.

ExSitu is interested in understanding the work practices of creative professionals, particularly artists, designers, and scientists, who push the limits of interactive technology. We follow a multi-disciplinary participatory design approach, working with both expert and non-expert users in diverse creative contexts. We also create situations that cause users to reflect deeply on their activities in situ and collaborate to articulate new design problems.

We identified diverse strategies for recording choreographic fragments and, influenced by the concept of *information substrates*, designed *Knotation* [19], a mobile pen-based tool where choreographers sketch representations of their choreographic ideas and make them interactive (Figure 5). Subsequent studies showed that *Knotation* supports both dance-then-record and record-then-dance strategies. Marianela Ciolfi Felice, supervised by Wendy Mackay and Sarah Fdili Alaoui, successfully defended her Ph.D. thesis *Supporting Expert Creative Practice* on this topic [32].

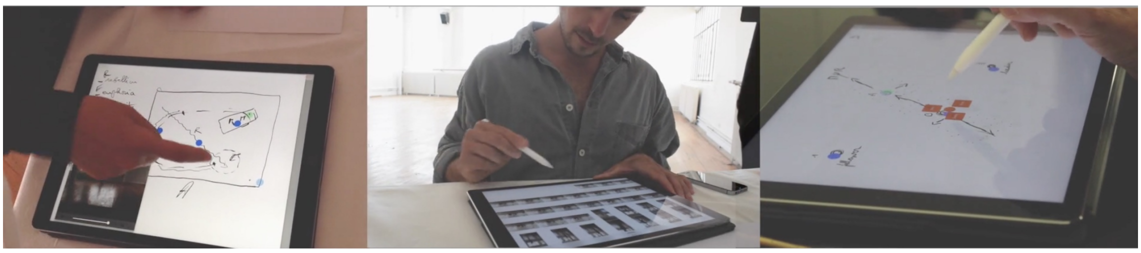


Figure 5. A choreographer uses Knotation to specify and interact with the spatial and temporal layout of a piece.

We are also developing a *Choreographer's Workbench*, a full-body interactive system that helps choreographers explore dance movements by linking previously recorded movement ideas and revealing their underlying relationships. The system emphasizes discoverability and appropriation of movement ideas, using feedforward to visualize movement characteristics. We studied how dancers learn complex expressive movements [23], and studied how variability during practice affects learning motor and timing skills [11]. We contributed to soma-based design, i.e. movement-based designs and design practices specifically engaging with aesthetics [13]. We also collaborated with Ircam on a tool that uses reinforcement learning to explore high-dimensional sound spaces [24]. Users enter likes and dislikes to guide navigation within the sound space, shifting from a parameter-based to a reward-based exploration strategy.

We also are interested in how makers transition between physical and digital designs. Makers often create both physical and digital prototypes to explore a design, taking advantage of the subtle feel of physical materials and the precision and power of digital models. We developed *ShapeMe* [25], a novel smart material that captures its own geometry as it is physically cut by an artist or designer. *ShapeMe* includes a software toolkit that lets its users generate customized, embeddable sensors that can accommodate various object shapes. As the designer works on a physical prototype, the toolkit streams the artist's physical changes to its digital counterpart in a 3D CAD environment (Figure 6). We used a rapid, inexpensive and simple-to-manufacture inkjet printing technique to create embedded sensors. We successfully created a linear predictive model of the sensors' lengths, and our empirical tests of *ShapeMe* showed an average accuracy of 2 to 3 mm. We further presented an application scenario for modeling multi-object constructions, such as architectural models, and 3D models consisting of multiple layers stacked one on top of each other.

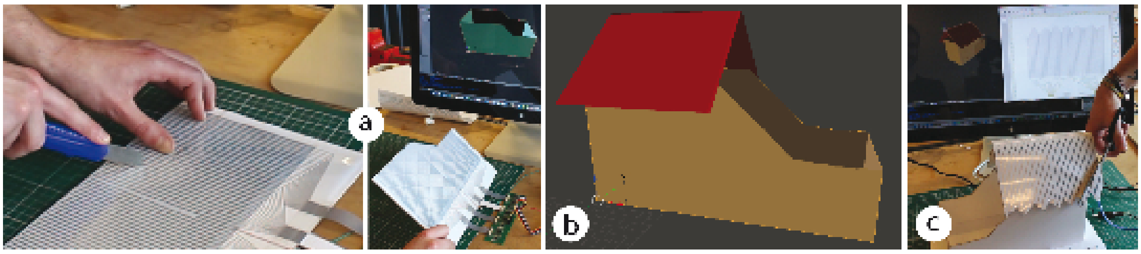


Figure 6. ShapeMe is a novel sensing technology that enables physical modeling with shape-aware material: (a) The maker cuts a foamcore piece to reshape the walls of a house model. The updated shape is captured by a grid of length-aware sensors and is communicated to 3D modeling software. (b) The makers digitally creates the pieces of the roof and then produces its physical model. (c) The maker explores variations of the roof by cutting its side with scissors, while its shape is continuously captured.

We also presented *Interactive Tangrami* [29], a method for prototyping interactive physical interfaces from functional paper-folded building blocks (Tangramis). *Interactive Tangrami* can contain various sensor input and visual output capabilities. Our digital design tool lets makers design the shape and interactive behavior of custom user interfaces. The software manages the communication with the paper-folded blocks and streams the interaction data via the Open Sound protocol (OSC) to an application prototyping environment, such as MaxMSP. The building blocks are fabricated digitally with a rapid and inexpensive ink-jet printing method. Our systems allows to prototype physical user interfaces within minutes and without knowledge of the underlying technologies. Finally, we continued our work with Saarland University, TU Berlin and MIT on digitally fabricated *directional screens* [15]. Michael Wessely, supervised by Theophanis Tsandilas and Wendy Mackay, successfully defended his Ph.D. thesis *Fabricating Malleable Interaction-Aware Material* [36] on these topics.

7.4. Collaboration

Participants: Cédric Fleury [correspondant], Michel Beaudouin-Lafon, Wendy Mackay, Carla Griggio, Yujiro Okuya.

ExSitu is interested in exploring new ways of supporting collaborative interaction and remote communication. We investigated how large interactive spaces such as wall-sized displays or immersive virtual reality systems can foster collaboration in both co-located and remote situations in the context of Digiscope (http://digiscope.fr/). We also conducted in-depth studies to better understand communication through social networks.

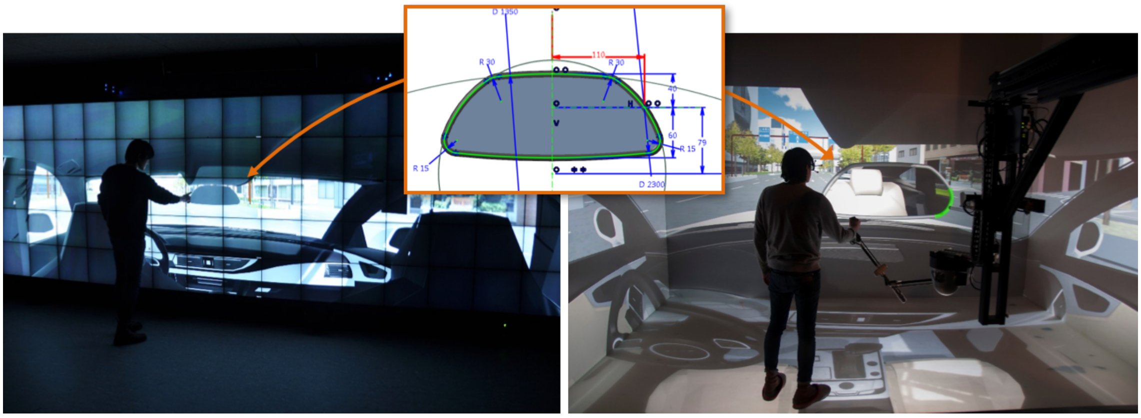


Figure 7. Collaborative CAD data modification between a wall-sized display (left) and a CAVE system (right).

Remote users can have significantly different display and interaction capabilities, such as a wall-size display v.s. an immersive CAVE. We started to explore how such asymmetric interaction capabilities provide interesting opportunities for new collaboration strategies. In particular, we developed a distributed architecture allowing the collaborative modifications of CAD data across heterogeneous platforms [27] and tested it between the EVE and WILDER platforms of Digiscope (CAVE vs. wall-sized touch display – Figure 7).

Remote collaboration across large interactive spaces also requires telepresence systems which support audio-video communication among users as they move in front of the display or inside of the immersive virtual reality system. We have added 3D audio to improve spatial awareness of remote users [26]: 3D audio lets us position a sound source for each remote participant at the virtual position occupied by this participant in the local space. When using video as well as audio, this lets us position the audio feed so that it is congruent with the position of the video feed.

Finally, we conducted an in-depth study of how users communicate via multiple social network apps that offer almost identical functionality. We studied how and why users distribute their contacts within their app ecosystem. We found that users appropriate the features and technical constraints of their apps to create idiosyncratic "communication places", each with its own recursively defined membership rules, perceived purposes, and emotional connotations. Users also shift the boundaries of their communication places to accommodate changes in their contacts' behavior, the dynamics of their relationships, and the restrictions of the technology. We argue that communication apps should support creating multiple "communication places" within the same app, relocating conversations across apps, and accessing functionality from other apps. Carla Griggio, supervised by Wendy Mackay, successfully defended her Ph.D. thesis *Designing for Ecosystems of Communication Apps* [33] on this topic.

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. MoveIT – Modeling the Speed/Accuracy Trade-Off of Human Aimed Movement with the Tools of Information Theory

Type: Ph.D. grant

Funding: DigiCosme Labex

Duration: 2015-2018

Coordinator: Olivier Rioul (Institut Mines Telecom)

Partners: Univ. Paris-Sud, Inria, CNRS, Institut Mines-Telecom

Inria contact: Michel Beaudouin-Lafon

Abstract: The goal of this project is to conduct fundamental studies of aimed movements based on information theory. The project studies the interaction phenomena involved in pointing, in order to discover novel, more effective pointing techniques. This project funds Wanyu Liu, a joint Ph.D. student between the COMELEC and VIA groups at Institut Mines Telecom and ExSitu. Wanyu defended her thesis in November 2018 [35] and received an Honorable Mention award for her CHI 2018 paper [22].

8.1.2. An Augmented-Reality System for Collaborative Physical Modeling and Design

Type: Equipment

Funding: STIC Paris-Saclay

Duration: 2017-2018

Coordinator: Theophanis Tsandilas

Partners: Univ. Paris-Sud, Inria

Inria contact: Theophanis Tsandilas

Abstract: The goal of the project is to develop an augmented-reality system to support collaboration over 3D models and enhance digital-fabrication approaches. It is a collaboration with the AVIZ group and provides funding (8k) for equipment.

8.1.3. Le Plateau des Recherches Infinies

Type: Equipment and subcontracting

Funding: Learning Center Paris-Saclay

Duration: 2017-2018

Coordinator: Michel Beaudouin-Lafon

Partners: Univ. Paris-Sud

Inria contact: Michel Beaudouin-Lafon

Abstract: The goal of this project (30k) is to create an interactive installation presenting the portraits of a hundred researchers from Université Paris-Saclay. It is a collaboration with portrait photographer Didier Goupy. The installation is designed to be exhibited in various sites of Université Paris-Saclay until it is permanently installed in the Learning Center of Université Paris-Saclay. This project supported Swati Swati, an intern, for two months over the summer. The project was presented at the Fête de la Science in October, 2018, and will be permanently exhibited in the future Learning Center of Université Paris-Saclay.

8.1.4. Virtual Reality for Interacting with Building Information Model at Paris-Saclay

Type: Equipment and human resources

Funding: STIC Paris-Saclay

Duration: 2018-2019

Coordinator: Jean-Marc Vézien (LIMSI-CNRS)

Partners: CNRS, Univ. Paris-Sud

Inria contact: Cédric Fleury

Abstract: The goal of this project is to develop interactive tools for BIM application in virtual reality using a user-centered design approach. The project will use as a case study the interior design of the *Learning Center* building on Paris-Saclay campus.

8.2. National Initiatives

8.2.1. ANR

ELEMENT: Enabling Learnability in Human Movement Interaction

Type: Equipment and human resources

Funding: ANR

Duration: 2019-2022

Coordinator: Baptiste Caramiaux, Sarah Fdili Alaoui, Wendy Mackay

Partners: IRCAM, LIMSI

Inria contact: Baptiste Caramiaux

Abstract: The goal of this project is to foster innovation in multimodal interaction, from non-verbal communication to interaction with digital media/content in creative applications, specifically by addressing two critical issues: the design of learnable gestures and movements; and the development of interaction models that adapt to a variety of user's expertise and facilitate human sensorimotor learning.

8.2.2. Investissements d'Avenir

8.2.2.1. Digiscope - Collaborative Interaction with Complex Data and Computation

Type: EQUIPEX (Equipement d'Excellence)

Duration: 2011-2019

Coordinator: Michel Beaudouin-Lafon

Partners: Université Paris-Saclay (coordinator), Université Paris-Sud, CNRS, CEA, Inria, Institut Mines-Telecom, CentraleSupelec, Université Versailles - Saint-Quentin, ENS Paris-Saclay, Maison de la Simulation

Overall budget: 22.5 Meuros, including 6.7 Meuros public funding from ANR

Abstract: The goal of the project is to create ten high-end interactive rooms interconnected by high-speed networks and audio-video facilities to support remote collaboration across interactive visualization environments. The equipment will be open to outside users and targets four main application areas: scientific discovery, product lifetime management, decision support for crisis management, and education and training. Digiscope includes the existing WILD room, and funded the WILDER room. ExSitu contributes its expertise in the design and evaluation of advanced interaction techniques and the development of distributed software architectures for interactive systems. All ten rooms and the telepresence network are operational. The project was successfully evaluated by an international jury in June, 2017.

8.3. European Initiatives

8.3.1. European Research Council (ERC)

8.3.1.1. Creating Human-Computer Partnerships

Program: ERC Advanced Grant

Project acronym: CREATIV

Project title: Creating Human-Computer Partnerships

Duration: June 2013 - May 2019

Coordinator: Wendy Mackay

Abstract: CREATIV explores how the concept of co-adaptation can revolutionize the design and use of interactive software. Co-adaptation is the parallel phenomenon in which users both adapt their behavior to the system's constraints, learning its power and idiosyncrasies, and appropriate the system for their own needs, often using it in ways unintended by the system designer. A key insight in designing for co-adaptation is that we can encapsulate interactions and treat them as first class objects, called interaction instruments This lets us focus on the specific characteristics of how human users express their intentions, both learning from and controlling the system. By making instruments co-adaptive, we can radically change how people use interactive systems, providing incrementally learnable paths that offer users greater expressive power and mastery of their technology. The initial goal of the CREATIV project is to fundamentally improve the learning and expressive capabilities of advanced users of creative software, offering significantly enhanced methods for expressing and exploring their ideas. The ultimate goal is to radically transform interactive systems for everyone by creating a powerful and flexible partnership between human users and interactive technology.

8.3.1.2. Unified Principles of Interaction

Program: ERC Advanced Grant

Project acronym: ONE

Project title: Unified Principles of Interaction

Duration: October 2016 - September 2020

Coordinator: Michel Beaudouin-Lafon

Abstract: The goal of ONE is to fundamentally re-think the basic principles and conceptual model of interactive systems to empower users by letting them appropriate their digital environment. The project addresses this challenge through three interleaved strands: empirical studies to better understand interaction in both the physical and digital worlds, theoretical work to create a conceptual model of interaction and interactive systems, and prototype development to test these principles and concepts in the lab and in the field. Drawing inspiration from physics, biology and psychology, the conceptual model combines *substrates* to manage digital information at various levels of abstraction and representation, *instruments* to manipulate substrates, and *environments* to organize substrates and instruments into digital workspaces.

8.3.2. Marie Skiodowska-Curie Actions

8.3.2.1. Enhancing Motion Interaction through Music Performance

Program: Marie Curie grant

Project acronym: MIM

Project title: Enhancing Motion Interaction

Duration: 2016 - 2018

Coordinator: Baptiste Caramiaux

Abstract: The goal of the project to enhance Human Motion–Computer Interaction by leveraging fa multidisciplinary approach across experimental psychology, music technology and computational modelling. Firstly, the project examines skilled activities, in particular music performance, in order to understand fundamental cognitive and psychological aspects of control and expression in human motion. The project involves computational models of motor control and expressive variations built from music performance data collected during psychophysical studies. Secondly, the project broaches the implementation of these models in Digital Musical Instruments (DMI), thus creating a new type of digital instrument based on sensorimotor learning mechanisms. The resulting DMI is then assessed through a user study in which elements of exploration and engagement will be tested over several sessions. Therefore, the project contributes to two main uncharted research areas. Firstly it contributes to the fundamental understanding of sensorimotor learning processes by considering complex human motion, specifically motion in music performance. Secondly, it represents an original application of computational modelling by modelling expressive musical gestures and transferring these models to interactive systems.

8.4. International Initiatives

8.4.1. Inria International Labs

Inria@SiliconValley

Associate Team involved in the International Lab:

8.4.1.1. DECibel

Title: Discover, Express, Create – Interaction Technologies For Creative Collaboration

International Partner (Institution - Laboratory - Researcher):

University of California Berkeley (United States) - Electrical and Computer Engineering, Center for Magnetic Resonance Research - Bjoern Hartmann

Start year: 2016

The DECibel associated team includes Inria's ExSitu and the CITRIS Connected Communities Initiative (CCI) at UC Berkeley. ExSitu explores extreme interaction, working with creative professionals and scientists who push the limits of technology to develop novel interactive technologies that offer new strategies for creative exploration. ExSitu's research activities include: developing underlying theory (co-adaptive instruments and substrates), conducting empirical studies (participatory design with creative professionals), and implementing interactive systems (creativity support tools). The CITRIS Connected Communities Initiative investigates collaborative discovery and design through new technologies that enhance education, creative work, and public engagement. It develops interactive tools, techniques and materials for the rapid design and prototyping of novel interactive products, expertise sharing among designers, and citizen science investigations. DECibel will combine the strengths of these two groups to to investigate novel tools and technologies that support Discovery, Expressivity, and Creativity.

8.5. International Research Visitors

8.5.1. Visits of International Scientists

Joanne McGrenere, Professor at the University of British Columbia, Canada and Inria Chair, visited for two months, to work with Wendy Mackay, Carla Griggio, Jessalyn Alvina, Yi Zhang and John MacCallum.

8.5.1.1. Internships

Janin Koch, Ph.D. student from Aalto University, Finland, visited for three months to work with Wendy Mackay.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. Member of the Organizing Committees

- ACM/ISMIR 2018, International Society for Music Information Retrieval Conference, Music and Technology co-chair: Wendy Mackay
- HAMAC International Research Workshop 2018, Human-Machine Collaboration in Embodied Interaction, main organiser: Baptiste Caramiaux
- UIST 2018, ACM Symposium on User Interface Software and Technology, Publicity Co-Chair: Michael Wessely
- UIST 2018, ACM Symposium on User Interface Software and Technology, Video Preview Co-Chair: Carla Griggio
- UIST 2018, ACM Symposium on User Interface Software and Technology, Doctoral Symposium jury: Wendy Mackay
- CHI 2018 workshop “Rethinking Interaction: From instrumental interaction to human-computer partnerships Toolkits”, organizers: Wendy Mackay, Michel Beaudouin-Lafon
- Interaction 2018, Program Co-Chair: Nolwenn Maudet

9.1.2. Scientific Events Selection

9.1.2.1. Chair of Conference Program Committees

- IHM 2018, Conférence Francophone d’Interaction Homme-Machine, Work-in-Progress (TeC) co-chair: Cédric Fleury

9.1.2.2. Member of the Conference Program Committees

- ACM CHI 2019, *ACM CHI Conference on Human Factors in Computing Systems*: Michel Beaudouin-Lafon, Wendy Mackay, Theophanis Tsandilas
- ACM UIST 2018, *ACM Symposium on User Interface Software and Technology*: Michel Beaudouin-Lafon, Wendy Mackay
- ACM UIST 2018 Doctoral Symposium: Wendy Mackay
- ACM UIST 2018 Best Paper Award Committee: Wendy Mackay
- ACM UIST 2018 Student Innovation Awards Committee
- IUI 2018, *ACM IUI Conference on Intelligent User Interfaces*: Wendy Mackay
- IUI 2019, *ACM IUI Conference on Intelligent User Interfaces*: Wendy Mackay
- MOCO 2018, *International Conference on Movement and Computing*: Sarah Fdili Alaoui
- 3DCVE 2018, *IEEE VR Workshop on Collaborative Virtual Environments*: Cédric Fleury
- ACM C&C 2018, *Creativity and Cognition*: Andrew Webb

9.1.2.3. Reviewer

- ACM CHI 2018-19, *ACM CHI Conference on Human Factors in Computing Systems*: Andrew Webb, Carla Griggio, Wanyu Liu, Sarah Fdili Alaoui, Cédric Fleury, Michael Wessely, Baptiste Caramiaux
- ACM UIST 2018, *ACM Symposium on User Interface Software and Technology*: Theophanis Tsandilas, Michael Wessely
- ACM DIS 2018, *ACM Conference on Designing Interactive Systems*: Nolwenn Maudet, Sarah Fdili Alaoui, Wendy Mackay
- ACM ISS 2018, *ACM International Conference on Interactive Surfaces and Spaces*: Ignacio Avellino, Philip Tchernavskij
- IEEE VR 2018-19, *Virtual Reality Conference*: Cédric Fleury
- Interaction 19: Carla Griggio, Germán Leiva
- ACM C&C 2018, *Creativity and Cognition*: Sarah Fdili Alaoui
- MOCO 2018, *Movement Computing*: Sarah Fdili Alaoui, Baptiste Caramiaux

- Mobile HCI '18 Late Breaking, *19th International Conference on Human-Computer Interaction with Mobile Devices and Services*: Germán Leiva, Philip Tchernavskij
- IHM 2018, *Conférence Francophone d'Interaction Homme-Machine*: Theophanis Tsandilas, Baptiste Caramiaux
- Salon des Refusés 2018, workshop at <PROGRAMMING> 2018, *The International Conference on the Art, Science, and Engineering of Programming*: Philip Tchernavskij

9.1.3. Journal

9.1.3.1. Member of Editorial Boards

- Editor for the Human-Computer Interaction area of the ACM Books Series (published with Morgan & Claypool Publishers): Michel Beaudouin-Lafon (2013-)
- CACM, *Communications of the ACM Web Editorial Board*, ACM: Wendy Mackay (2008-)
- ACM, *ACM New Publications Board*, ACM: Wendy Mackay (2015-)
- TOCHI, *Transactions on Computer Human Interaction*, ACM: Michel Beaudouin-Lafon (2009-), Wendy Mackay (2016-)
- PloS ONE: Baptiste Caramiaux (2018-)
- JIPS, *Journal d'Interaction Personne-Système*, AFIHM: Michel Beaudouin-Lafon (2009-)

9.1.3.2. Reviewer - Reviewing Activities

- TOCHI, *Transactions on Computer Human Interaction*, ACM: Baptiste Caramiaux
- CG&A, *Computer Graphics and Applications*, IEEE: Cédric Fleury
- Frontiers in *Robotics and AI*, Virtual Environments section: Cédric Fleury

9.1.4. Invited Talks

- Univ. of Calfornia San Diego (UCSD), *Towards Unified Principles of Interaction*, March 2018: Michel Beaudouin-Lafon
- Univ. of Calfornia San Diego (UCSD), *Human-Computer Partnerships*, March 2018: Wendy Mackay
- ETH Zürich, *Towards Unified Principles of Interaction*, 11 April 2018: Michel Beaudouin-Lafon
- Université de Zürich, *Human-Computer Partnerships*, 12 April 2018: Wendy Mackay
- Creativity Across Disciplines Symposium, Aarhus University, *Towards a Unified Theory of Interaction*, 5 April 2018: Wendy Mackay
- Workshop on Human-Computer Partnerships, ACM CHI 2018, *Human-in-the-loop? or Human-Computer Partnerships?*, May 2018: Wendy Mackay
- SIG: Transparent Statistics Guidelines, Panel Discussion, ACM CHI 2018, May 2018: Wendy Mackay
- SIFED 2018 : Symposium International Francophone sur l'Ecrit et le Document, *Interactions à base d'esquisses sur papier et ordinateur*, 1 June 2018, Tours, France: Theophanis Tsandilas
- DATAIA-JST International Symposium on Data Science and AI, *AI and HCI: Towards Human-Computer Partnerships*, Paris, 11 July 2018: Michel Beaudouin-Lafon
- Science & Research Panel, *Interactive Materials with Printed Electronics*, 2018 Fab14 Paris, 13 July 2018: Micheal Wessely
- Rencontres Jeunes Chercheurs, RJIHM, *Experiment Design in HCI*, July: Wendy Mackay
- Workshop on Human-Computer Collaboration in Embodied Interaction (HAMAC), *Creating Human-Computer Partnerships*, IRCAM, Paris, 8 July 2018: Michel Beaudouin-Lafon
- Inria seminar Unithé ou Café, *Machine Learning and Teaching for Interaction Design*, 13 Septembre 2018: Baptiste Caramiaux

- Visual Computing Conference, Stuttgart, *Human-Computer Partnerships*, 8 October 2018: Wendy Mackay
- ECSS 2018, *European Computer Science Summit, Informatics Europe*, Gothenberg, 11 October 2018: Wendy Mackay
- GOTO Copenhagen, Keynote address, *Human-Computer Partnerships*, November 2018: Wendy Mackay
- Université de Poitiers, Techné Séminaire 1 *Technologie Numérique pour l'Education*, *Designing Interactive Systems*, 15 November 2018: Wendy Mackay
- Université de Poitiers, Techné Séminaire 2 *Technologie Numérique pour l'Education*, *Exploratory Experiment Design*: 15 November 2018: Wendy Mackay

9.1.5. Scientific Expertise

International

- NSERC Evaluation: Wendy Mackay, external expert (2018)
- ACM SIGCHI "Lifetime Service Award" committee member: Michel Beaudouin-Lafon

National

- Agence Nationale de la Recherche (ANR), Appel à projets jeunes chercheuses et jeunes chercheurs: Theophanis Tsandilas, reviewer
- ERC Generator program, Université de Lille: Michel Beaudouin-Lafon, external expert
- CNRS INS2I "Cellule ERC": Michel Beaudouin-Lafon, member

9.1.6. Research Administration

Information Science and Technology (STIC) Department, Université Paris-Saclay: Michel Beauouin-Lafon (chair since June 2018), Wendy Mackay (member)

Research division, Université Paris-Saclay: Michel Beaudouin-Lafon (advisor for Digital Sciences since June 2018)

Digiteo RTRA research network, Université Paris-Saclay: Michel Beauouin-Lafon (director since June 2018)

STIC Doctoral School, Université Paris-Saclay: Michel Beaudouin-Lafon (adjunct director of pole 3 until May 2018)

CNRS INS2I, "Conseil Scientifique de l'Institut": Michel Beaudouin-Lafon (member)

Telecom ParisTech, "Comité de la recherche": Michel Beaudouin-Lafon (member)

Pôle Systematic, Working group on Information Systems: Michel Beaudouin-Lafon (member of steering committee)

Computer Science Department, Université Paris-Sud: Michel Beaudouin-Lafon (vice-President for research)

"Conseil de Laboratoire", LRI: Wendy Mackay, Cédric Fleury (members)

"Conseil Scientifique", LRI: Michel Beaudouin-Lafon (member)

CCSU, "Commission Consultative de Spécialistes de l'Université", Université Paris-Sud: Michel Beaudouin-Lafon, Wendy Mackay (members)

"Commission Locaux", LRI: Theophanis Tsandilas (member)

"Commission Scientifique", Inria: Theophanis Tsandilas (member), since March 2017

"Comité de sélection, Professeur", Université de Bordeaux: Michel Beaudouin-Lafon (president)

"Comité de sélection, Maître de Conférence", Université Paris-Sud: Wendy Mackay

"Tenure review committee", University of California, Irvine: Wendy Mackay

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence : Sarah Fdili Alaoui, *Programmation des interfaces interactives avancées*, 22.5h, L3, Univ. Paris-Sud

International Masters: Theophanis Tsandilas, *Probabilities and Statistics*, 32h, M1, Univ. Paris-Saclay

HCID Masters: Sarah Fdili Alaoui, *Business Development Labs*, 30h, M1, Univ. Paris-Sud

HCID Masters: Sarah Fdili Alaoui, *Innovation & Entrepreneurship thesis*, 3h, M2, Univ. Paris-Sud

HCID Masters: Sarah Fdili Alaoui, *Design Project*, 36h, M1 et M2, Univ. Paris-Sud

HCID Masters: Michel Beaudouin-Lafon, Wendy Mackay, *Fundamentals of Situated Interaction*, 21 hrs, M1/M2, Univ. Paris-Sud

Interaction & HCID Masters: Sarah Fdili Alaoui, *Stage en entreprise*, 2h, M2, Univ. Paris-Sud

Interaction & HCID Masters: Sarah Fdili Alaoui, *Creative Design*, 27h, M1 et M2, Univ. Paris-Sud

Interaction & HCID Masters: Sarah Fdili Alaoui, *Digital Fabrication*, 13,5h, M1 et M2, Univ. Paris-Sud

Interaction & HCID Masters: Michel Beaudouin-Lafon, *Fundamentals of Human-Computer Interaction*, 21 hrs, M1/M2, Univ. Paris-Sud

Interaction & HCID Masters: Michel Beaudouin-Lafon & Cédric Fleury, *Groupware and Collaborative Interaction*, 31.5 hrs, M1/M2, Univ. Paris-Sud

Interaction & HCID Masters: Wendy Mackay, *Career Seminar* 6 hrs, M2, Univ. Paris-Sud

Interaction & HCID Masters: Wendy Mackay, *Design of Interactive Systems*, 21 hrs, M1/M2, Univ. Paris-Sud

Interaction & HCID Masters: Wendy Mackay, *Advanced Design of Interactive Systems*, 21 hrs, M1/M2, Univ. Paris-Sud

Interaction & HCID Masters: Baptiste Caramiaux, *Gestural and Mobile Interaction*, 24 hrs, M1/M2, Univ. Paris-Sud

Polytech: Sarah Fdili Alaoui, *Graphisme et Visualisation*, 18h, “Apprentis” 5th year, Univ. Paris-Sud

Polytech: Cédric Fleury, *Projet Java-Graphique-IHM*, 24 hrs, 3st year, Univ. Paris-Sud

Polytech: Cédric Fleury, *Interaction Homme-Machine*, 18 hrs, 3st year, Univ. Paris-Sud

Polytech: Cédric Fleury, *Option Réalité Virtuelle*, 56 hrs, 5th year, Univ. Paris-Sud

Polytech: Cédric Fleury, *Réalité Virtuelle et Interaction*, 48 hrs, “Apprentis” 5th year, Univ. Paris-Sud

9.2.2. Supervision

PhD students

PhD: Wanyu Liu, *Modeling the speed-accuracy trade-off of pointing tasks using the tools of information theory*, 22 November 2018. Advisors: Olivier Rioul (Institut Mines Telecom) & Michel Beaudouin-Lafon

PhD: Carla Griggio, *Interactive Human-Machine Learning*, Université Paris-Saclay, 11 December 2018. Advisor: Wendy Mackay

PhD: Germán Leiva, *Interaction-driven Software Development*, Université Paris-Saclay, 12 December 2018. Advisors: Wendy Mackay & Michel Beaudouin-Lafon

PhD: Michael Wessely, *Sketching and Physical Prototyping for Creative Fabrication Design*, Université Paris-Saclay, 13 December 2018. Advisors: Theophanis Tsandilas & Wendy Mackay

PhD: Marianela Ciolfi Felice, *Substrates and Co-adaptive Instruments for Creativity*, Université Paris-Saclay, 14 December 2018. Advisors: Wendy Mackay & Sarah Fdili Alaoui

PhD in progress: Yujiro Okuya, *Sensorimotor interface for Collaborative Virtual Environments based on heterogeneous interactive devices: application to industrial design*, October 2015. Advisors: Patrick Bourdot (LIMSI-CNRS) & Cédric Fleury

PhD in progress: Stacy (Shu-Yuan) Hsueh, *Embodied design for Human-Computer Co-creation*, November 2015. Advisors: Wendy Mackay & Sarah Fdili Alaoui

PhD in progress: Philip Tchernavskij, *Towards Unified Principles of Interaction*, October 2016. Advisor: Michel Beaudouin-Lafon

PhD in progress: Jean-Philippe Rivière, *Embodied Design for Human-Computer Partnership in Learning Contexts*, October 2017. Advisors: Wendy Mackay, Sarah Fdili Alaoui & Baptiste Caramiaux

PhD in progress: Yiran Zhang, *Telepresence for remote and heterogeneous Collaborative Virtual Environments*, October 2017. Advisors: Patrick Bourdot (LIMSI-CNRS) & Cédric Fleury

PhD in progress: Téo Sanchez, *Co-Learning in Interactive Systems*, September 2018. Advisors: Baptiste Caramiaux & Wendy Mackay

PhD in progress: Elizabeth Walton, *Inclusive Design in Embodied Interaction*, November 2018. Advisors: Baptiste Caramiaux, Sarah Fdili Alaoui & Wendy Mackay

PhD in progress: Miguel Renom, *Theoretical bases of human tool use in digital environments*, October 2018. Advisors: Michel Beaudouin-Lafon & Baptiste Caramiaux

PhD in progress: Dimitrios Christaras Papageorgiou, *Design, implementation and evaluation of a conceptual model of interactive substrates*, October 2018. Advisor: Michel Beaudouin-Lafon

PhD in progress: Han Han, *Participatory design of digital environments based on interaction substrates*, October 2018. Advisor: Michel Beaudouin-Lafon

PhD in progress: Viktor Gustafsson, *Co-adaptive Instruments fo Game Design*, October 2018. Advisor: Wendy Mackay

PhD in progress: Yi Zhang, *Generative Design using Instrumental Interaction, Substrates and Co-adaptive Systems*, October 2018. Advisor: Wendy Mackay

Masters students

Téo Sanchez, ENS, Sorbonne Université, "Gesture Interaction with Active Machine Learners": Baptiste Caramiaux

Miguel Renom, "Transfer of tool-based skills from physical to digital tools": Michel Beaudouin-Lafon

Dimitrios Christaras Papageorgiou, "Prototype of an environment based on information substrates and digital instruments": Michel Beaudouin-Lafon

Tong Xue, "Analysis of information substrates and digital instruments in existing interactive software": Michel Beaudouin-Lafon

Viktor Gustafsson, "Narrative Substrates for Interactive Video Games": Wendy Mackay

Yi Zhang, "Situated Breakdowns in Video-Mediated Communication": Wendy Mackay

Wuji Geng, "Interaction Museum Design": Wendy Mackay

Krishnan Chandran, "Collocated Collaboration over Augmented Reality Models": Theophanis Tsandilas and Cédric Fleury

9.2.3. Juries

PhD theses

Emaline Brulé, Télécom ParisTech (advisor: Annie Gentes): Wendy Mackay, reporter

Julien Gori, Télécom ParisTech (advisor: Olivier Rioul): Michel Beaudouin-Lafon, invited

Benjamen Bressolette, Centrale Marseille (advisor: Richard Kronland-Martinet): Michel Beaudouin-Lafon, examiner

Habilitations

- Marcos Serrano, Université de Toulouse, November 2018: Michel Beaudouin-Lafon, reviewer
- Eric Lecolinet, Télécom ParisTech, December 2018: Michel Beaudouin-Lafon, examiner

9.3. Popularization

9.3.1. Articles and contents

- Radio show “Les P’tits Bateaux” on France Inter public radio: Michel Beaudouin-Lafon (interviews in response to children questions on 10 June 2018, 16 September 2018, 7 October 2018)
- Panel at “*La Fabrique de la Danse*”, Centre 104, Paris, 5 Avril 2018: Baptiste Caramiaux (panel member)

9.3.2. Interventions

- Le Plateau des Recherches Infinies: art-science project with photographer Didier Goupy presenting the portraits of 100 researchers of Université Paris-Saclay. Fête de la Science, Gif-sur-Yvette, October 2018
- Workshop on “Rapid Prototyping with Interactive Materials", Fab14, Toulouse, July 2018: Michael Wessely
- Workshop “*Relational Listening*”, CLOUD/Danslab, Lake Studios, Berlin, 22 July 2018: John MacCallum (co-facilitator).

10. Bibliography

Major publications by the team in recent years

[1] S. F. ALAOUI, J. FRANÇOISE, T. SCHIPHORST, K. STUDD, F. BEVILACQUA.*Seeing, Sensing and Recognizing Laban Movement Qualities*, in "ACM Conference on Human Factors in Computing Systems (CHI)", Denver, United States, May 2017 [*DOI :* 10.1145/3025453.3025530], https://hal.inria.fr/hal-01663132

[2] J. ALVINA, J. MALLOCH, W. MACKAY.*Expressive Keyboards: Enriching Gesture-Typing on Mobile Devices*, in "Proceedings of the 29th ACM Symposium on User Interface Software and Technology (UIST 2016)", Tokyo, Japan, ACM, October 2016, p. 583 - 593 [*DOI :* 10.1145/2984511.2984560], https://hal.inria.fr/hal-01437054

[3] I. AVELLINO, C. FLEURY, W. MACKAY, M. BEAUDOUIN-LAFON.*CamRay: Camera Arrays Support Remote Collaboration on Wall-Sized Displays*, in "Proceedings of the CHI Conference on Human Factors in Computing Systems", Denver, United States, CHI ’17, ACM, May 2017, p. 6718 - 6729 [*DOI :* 10.1145/3025453.3025604], https://hal.archives-ouvertes.fr/hal-01544645

[4] M. C. CIOLFI FELICE, N. MAUDET, W. E. MACKAY, M. BEAUDOUIN-LAFON.*Beyond Snapping: Persistent, Tweakable Alignment and Distribution with StickyLines*, in "UIST ’16 Proceedings of the 29th Annual Symposium on User Interface Software and Technology", Tokyo, Japan, Proceedings of the 29th Annual Symposium on User Interface Software and Technology, October 2016 [*DOI :* 10.1145/2984511.2984577], https://hal.archives-ouvertes.fr/hal-01410171

[5] C. KLOKMOSE, J. EAGAN, S. BAADER, W. MACKAY, M. BEAUDOUIN-LAFON.*Webstrates: Shareable Dynamic Media*, in "ACM Symposium on User Interface Software and Technology (UIST)", Charlotte, United States, UIST ’15 Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology, ACM, November 2015, p. 280-290 [*DOI :* 10.1145/2807442.2807446], https://hal.archives-ouvertes.fr/hal-01242672

[6] C. LIU, O. CHAPUIS, M. BEAUDOUIN-LAFON, E. LECOLINET.*Shared Interaction on a Wall-Sized Display in a Data Manipulation Task*, in "Proceedings of the 34th international conference on Human factors in computing systems", San Jose, United States, CHI '16, SigCHI and ACM, May 2016, p. 1-12 [*DOI :* 10.1145/2858036.2858039], https://hal.inria.fr/hal-01275535

[7] W. LIU, R. LUCAS D'OLIVEIRA, M. BEAUDOUIN-LAFON, O. RIOUL.*BIGnav: Bayesian Information Gain for Guiding Multiscale Navigation*, in "ACM CHI 2017 - International conference of Human-Computer Interaction", Denver, United States, May 2017, p. 5869-5880 [*DOI :* 10.1145/3025453.3025524], https://hal.inria.fr/hal-01677122

[8] M. NANCEL, E. PIETRIGA, O. CHAPUIS, M. BEAUDOUIN-LAFON.*Mid-air Pointing on Ultra-Walls*, in "ACM Transactions on Computer-Human Interaction", August 2015, vol. 22, n^o 5, p. 21:1–21:62 [*DOI :* 10.1145/2766448], https://hal.inria.fr/hal-01184544

[9] T. TSANDILAS, A. BEZERIANOS, T. JACOB.*SketchSliders: Sketching Widgets for Visual Exploration on Wall Displays*, in "Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems", Seoul, South Korea, ACM, April 2015, p. 3255-3264 [*DOI :* 10.1145/2702123.2702129], https://hal.archives-ouvertes.fr/hal-01144312

[10] M. WESSELY, T. TSANDILAS, W. E. MACKAY.*Stretchis: Fabricating Highly Stretchable User Interfaces*, in "ACM Symposium on User Interface Software and Technology (UIST)", Tokyo, Japan, October 2016, p. 697-704 [*DOI :* 10.1145/2984511.2984521], https://hal.archives-ouvertes.fr/hal-01412289

Publications of the year

Articles in International Peer-Reviewed Journal

[11] B. CARAMIAUX, F. BEVILACQUA, M. M. WANDERLEY, C. PALMER.*Dissociable effects of practice variability on learning motor and timing skills*, in "PLoS ONE", 2018, vol. 13, n^o 3, e0193580 [*DOI :* 10.1371/JOURNAL.PONE.0193580], https://hal.archives-ouvertes.fr/hal-01726186

[12] A. GOGOLOU, T. TSANDILAS, T. PALPANAS, A. BEZERIANOS.*Comparing Similarity Perception in Time Series Visualizations*, in "IEEE Transactions on Visualization and Computer Graphics", October 2018, 11, https://hal.inria.fr/hal-01845008

[13] K. HÖÖK, B. CARAMIAUX, C. ERKUT, J. FORLIZZI, N. HAJINEJAD, M. HALLER, C. C. M. HUMMELS, K. ISBISTER, M. JONSSON, G. KHUT, L. LOKE, D. LOTTRIDGE, P. MARTI, E. MELCER, F. F. MÜLLER, M. GRAVES PETERSEN, T. SCHIPHORST, E. M. SEGURA, A. STÅHL, D. SVANAES, J. THOLANDER, H. TOBIASSON.*Embracing First-Person Perspectives in Soma-Based Design*, in "Informatics", February 2018, vol. 5, n^o 1 [*DOI :* 10.3390/INFORMATICS5010008], https://hal.archives-ouvertes.fr/hal-01699005

[14] Y. OKUYA, N. LADEVEZE, C. FLEURY, P. BOURDOT.*ShapeGuide: Shape-Based 3D Interaction for Parameter Modification of Native CAD Data*, in "Frontiers in Robotics and AI", November 2018, vol. 5 [*DOI :* 10.3389/FROBT.2018.00118], https://hal.inria.fr/hal-01958226

[15] M. PIOVARČI, M. WESSELY, M. JAGIELSKI, M. ALEXA, W. MATUSIK, P. DIDYK.*Design and analysis of directional front projection screens*, in "Computers and Graphics", August 2018, vol. 74, p. 213 - 224 [*DOI :* 10.1016/J.CAG.2018.05.010], https://hal.archives-ouvertes.fr/hal-01848680

[16] T. TSANDILAS.*Fallacies of Agreement: A Critical Review of Consensus Assessment Methods for Gesture Elicitation*, in "ACM Transactions on Computer-Human Interaction", June 2018, vol. 25, n^{o} 3, p. 1-49 [*DOI :* 10.1145/3182168], https://hal.archives-ouvertes.fr/hal-01788775

International Conferences with Proceedings

[17] A. BASMAN, P. TCHERNAVSKIJ, S. BATES, M. BEAUDOUIN-LAFON.*An anatomy of interaction: co-occurrences and entanglements*, in "Programming'18 Companion - Conference Companion of the 2nd International Conference on Art, Science, and Engineering of Programming", Nice, France, ACM Press, April 2018, p. 188-196 [*DOI :* 10.1145/3191697.3214328], https://hal.archives-ouvertes.fr/hal-01854418

[18] M. BEAUDOUIN-LAFON, W. MACKAY.*Rethinking Interaction: From Instrumental Interaction to Human-Computer Partnerships*, in "Extended Abstracts of the 2018 CHI Conference", Montreal QC, Canada, ACM Press, April 2018, CHI 2018 Workshop [*DOI :* 10.1145/3170427.3170635], https://hal.archives-ouvertes.fr/hal-01788023

[19] M. CIOLFI FELICE, S. FDILI ALAOUI, W. MACKAY.*Knotation: Exploring and Documenting Choreographic Processes*, in "Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems", Montreal QC, Canada, April 2018 [*DOI :* 10.1145/3173574.3174022], https://hal.inria.fr/hal-01817569

[20] J. GORI, O. RIOUL, Y. GUIARD, M. BEAUDOUIN-LAFON.*The Perils of Confounding Factors: How Fitts' Law Experiments can Lead to False Conclusions*, in "CHI' 18", Montréal, Canada, April 2018 [*DOI :* 10.1145/3173574.3173770], https://hal.archives-ouvertes.fr/hal-01690122

[21] G. LEIVA, M. BEAUDOUIN-LAFON.*Montage: A Video Prototyping System to Reduce Re-Shooting and Increase Re-Usability*, in "UIST '18", Berlin, Germany, October 2018 [*DOI :* 10.1145/3242587.3242613], https://hal.archives-ouvertes.fr/hal-01966544

[22] W. LIU, O. RIOUL, J. MCGRENERE, W. MACKAY, M. BEAUDOUIN-LAFON.*BIGFile: Bayesian Information Gain for Fast File Retrieval*, in "Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)", Montreal QC, Canada, ACM Press, April 2018 [*DOI :* 10.1145/3173574.3173959], https://hal.archives-ouvertes.fr/hal-01791754

[23] J.-P. RIVIÈRE, S. FDILI ALAOUI, B. CARAMIAUX, W. MACKAY.*How Do Dancers Learn To Dance?: A first-person perspective of dance acquisition by expert contemporary dancers*, in "MOCO 2018 - 5th International Conference on Movement and Computing", Gênes, Italy, June 2018, p. 1-8 [*DOI :* 10.1145/3212721.3212723], https://hal.archives-ouvertes.fr/hal-01849604

[24] H. SCURTO, F. BEVILACQUA, B. CARAMIAUX.*Perceiving Agent Collaborative Sonic Exploration In Interactive Reinforcement Learning*, in "Proceedings of the 15th Sound and Music Computing Conference (SMC 2018)", Limassol, Cyprus, July 2018, https://hal.archives-ouvertes.fr/hal-01849074

[25] M. WESSELY, T. TSANDILAS, W. MACKAY.*Shape-Aware Material: Interactive Fabrication with ShapeMe*, in "UIST 2018 - 31st ACM Symposium on User Interface Software and Technology", Berlin, Germany, ACM Press, October 2018, vol. 18, p. 127-139 [*DOI :* 10.1145/3242587.3242619], https://hal.archives-ouvertes.fr/hal-01872776

Conferences without Proceedings

[26] L. FYFE, O. GLADIN, C. FLEURY, M. BEAUDOUIN-LAFON.*Combining Web Audio Streaming, Motion Capture, and Binaural Audio in a Telepresence System*, in "Web Audio Conference", Berlin, Germany, September 2018, https://hal.inria.fr/hal-01957843

[27] Y. OKUYA, N. LADEVEZE, O. GLADIN, C. FLEURY, P. BOURDOT.*Distributed Architecture for Remote Collaborative Modification of Parametric CAD Data*, in "IEEE VR International Workshop on 3D Collaborative Virtual Environments (3DCVE 2018)", Reutlingen, Germany, March 2018, https://hal.inria.fr/hal-01849289

[28] P. TCHERNAVSKIJ.*Critique of 'files as directories: some thoughts on accessing structured data within files' (1)*, in "Conference Companion of the 2nd International Conference on Art, Science, and Engineering of Programming", Nice, France, ACM Press, April 2018 [*DOI :* 10.1145/3191697.3214324], https://hal.archives-ouvertes.fr/hal-01854295

[29] M. WESSELY, N. MORENKO, J. STEIMLE, M. SCHMITZ.*Interactive Tangrami: Rapid Prototyping with Modular, Paper-folded Electronics*, in "Symposium on User Interface Software and Technology", Berlin, Germany, October 2018, p. 143-145 [*DOI :* 10.1145/3266037.3271630], https://hal.archives-ouvertes.fr/hal-01959161

Research Reports

[30] C. BARILLOT, I. FIJALKOW, I. QUEINNEC, F. THEOLEYRE, H. TOUZET, M. BEAUDOUIN-LAFON, M.-P. CANI, F. CHAUMETTE, G. CONREUR, V. CORTIER, C. GAVOILLE, G. –. GOGNIAT, J. GOSSA, A. HERZIG, J. KRIVINE, P. LAMARRE, F. LAMNABHI-LAGARRIGUE, A.-C. LETOURNEL, L. SEINTURIER, I. TELLIER, S. TORRES, C. TRUCHET, A. TSOUKIAS, M. VERLEYSEN.*Prospective Report of the scientific council of the Information Sciences and their interactions Institute of CNRS*, CNRS, December 2018, https://hal.archives-ouvertes.fr/hal-01956087

[31] A. GOGOLOU, T. TSANDILAS, T. PALPANAS, A. BEZERIANOS.*Comparing Time Series Similarity Perception under Different Color Interpolations*, Inria, June 2018, n^o RR-9189, https://hal.inria.fr/hal-01844994

References in notes

[32] M. C. FELICE.*Supporting Expert Creative Practice*, Université Paris-Saclay, Orsay, France, December 2018

[33] C. GRIGGIO.*Designing for Ecosystems of Communication Apps*, Université Paris-Saclay, Orsay, France, December 2018

[34] G. LEIVA.*Interactive Prototyping of Interactions: From Throwaway Prototypes to Takeaway Prototyping*, Université Paris-Saclay, Orsay, France, December 2018

[35] W. LIU.*Information theory as a unified tool for understanding and designing human-computer interaction*, Université Paris-Saclay, Orsay, France, November 2018

[36] M. WESSELY.*Fabricating Malleable Interaction-Aware Material*, Université Paris-Saclay, Orsay, France, December 2018

Team GALEN-POST

Organ Modeling through Extraction, Representation and Understanding of Medical Image Content

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

RESEARCH CENTER
Saclay - Île-de-France

THEME
Computational Neuroscience and Medicine

Table of contents

Team GALEN-POST

Creation of the Team: 2018 January 01

Keywords:

Computer Science and Digital Science:

A5.3. - Image processing and analysis
A5.4. - Computer vision
A5.9. - Signal processing
A8. - Mathematics of computing
A8.2. - Optimization
A8.7. - Graph theory
A9. - Artificial intelligence
A9.2. - Machine learning
A9.3. - Signal analysis

Other Research Topics and Application Domains:

B1. - Life sciences
B2. - Health
B2.6. - Biological and medical imaging

1. Team, Visitors, External Collaborators

Faculty Members

Jean-Christophe Pesquet [Team leader, Centrale-Supélec, Professor, HDR]
Hugues Talbot [Centrale-Supélec, Professor, from May 2018, HDR]
Fragkiskos Malliaros [Centrale-Supélec, Associate Professor, until Sep 2018]
Edouard Oyallon [Centrale-Supélec, Assistant Professor]
Emilie Chouzenoux [Univ. Paris-Est Marne La Vallée, Associate Professor, HDR]
Marc Castella [Télécom SudParis, Associate Professor]

External Collaborator

Maria Papadomanolaki [Centrale-Supélec]

Technical Staff

Sagar Verma [Centrale-Supélec, from Jul 2018]

PhD Students

Enzo Battistella [Centrale-Supélec, from Sep 2018]
Eugène Belilovsky [Centrale-Supélec, until Mar 2018]
Abdulkadir Celikkanat [Centrale-Supélec]
Siddhartha Chandra [Centrale-Supélec, until Mar 2018]
Guillaume Chassagnon [Centrale-Supélec]
Marie-Caroline Corbineau [Centrale-Supélec]
Theo Estienne [Centrale-Supélec]
Riza Guler [Inria, until Oct 2018]
Yunshi Huang [Univ. Paris East]
Hariprasad Kannan [Centrale-Supélec, until Jun 2018]
Stefan Kinauer [Centrale-Supélec, until Mar 2018]
Huu Dien Khue Le [Centrale-Supélec, until Mar 2018]

Marvin Lerousseau [Centrale-Supélec, from Oct 2018]
Arthur Marmin [Centrale-Supélec]
Mihir Sahasrabudhe [Centrale-Supélec]
Maissa Sghaier [Centrale-Supélec]
Roger Sun [Centrale-Supélec]

Post-Doctoral Fellow

Maria Vakalopoulou [Centrale-Supélec, until Nov 2018]

Administrative Assistants

Alexandra Merlin [Inria]
Jana Dutrey [Centrale-Supélec]

2. Overall Objectives

2.1. GALEN-POST Centrale-Supélec

Computational vision is one of the most challenging research domains in engineering sciences. The aim is to reproduce human visual perception through intelligent processing of visual data. The application domains span from computer aided diagnosis to industrial automation & robotics. The most common mathematical formulation to address such a challenge is through mathematical modeling. In such a context, first the solution of the desired vision task is expressed in the form of a parameterized mathematical model. Given such a model, the next task consists of associating the model parameters with the available observations, which is often called the model-to-data association. The aim of this task is to determine the impact of a parameter choice to the observations and eventually maximize/minimize the adequacy of these parameters with the visual observations. In simple words, the better the solution is, the better it will be able to express and fit the data. This is often achieved through the definition of an objective function on the parametric space of the model. Last, but not least given the definition of the objective function, visual perception is addressed through its optimization with respect to the model parameters. To summarize, computation visual perception involves three aspects, a task-specific definition of a parametric model, a data-specific association of this model with the available observations and last the optimization of the model parameters given the objective and the observations.

Such a chain processing inherits important shortcomings. The curse of dimensionality is often used to express the importance of the model complexity. In simple words, the higher the complexity of the model is, the better its expressive power will be with counter effect the increase of the difficulty of the inference process. Non-linearity is another issue to be addressed which simply states that the association between the model and the data is a (highly) non-linear function and therefore direct inference is almost infeasible. The impact of this aspect is enforced from the curse of non-convexity that characterizes the objective function. Often it lives in high-dimensional spaces and is ill posed making exact inference problematic (in many cases not possible) and computationally expensive. Last, but not least modularity and scalability is another important concern to be addressed in the context of computational vision. The use of task-specific modeling and algorithmic solutions make their portability infeasible and therefore transfer of knowledge from one task to another is not straightforward while the methods do not always scale well with respect either to the dimensionality of the representation or the data.

GALEN aims at proposing innovative techniques towards automatic structuring, interpretation and longitudinal modeling of visual data. In order to address these fundamental problems of computational perception, GALEN investigates the use of discrete models of varying complexity. These methods exhibit an important number of strengths such as their ability to be modular with respect to the input measurements (clinical data), the nature of the model (certain constraints are imposed from computational perspective in terms of the level of interactions), and the model-to-data association while being computational efficient.

3. Research Program

3.1. Shape, Grouping and Recognition

A general framework for the fundamental problems of image segmentation, object recognition and scene analysis is the interpretation of an image in terms of a set of symbols and relations among them. Abstractly stated, image interpretation amounts to mapping an observed image, X to a set of symbols Y. Of particular interest are the symbols Y^* that *optimally explain the underlying image*, as measured by a scoring function s that aims at distinguishing correct (consistent with human labellings) from incorrect interpretations:

$$Y^* = \mathrm{argmax}_Y s(X, Y) \tag{1}$$

Applying this framework requires (a) identifying which symbols and relations to use (b) learning a scoring function s from training data and (c) optimizing over Y in Eq.1.

One of the main themes of our work is the development of methods that jointly address (a,b,c) in a shape-grouping framework in order to reliably extract, describe, model and detect shape information from natural and medical images. A principal motivation for using a shape-based framework is the understanding that shape- and more generally, grouping- based representations can go all the way from image features to objects. Regarding aspect (a), image representation, we cater for the extraction of image features that respect the shape properties of image structures. Such features are typically constructed to be purely geometric (e.g. boundaries, symmetry axes, image segments), or appearance-based, such as image descriptors. The use of machine learning has been shown to facilitate the robust and efficient extraction of such features, while the grouping of local evidence is known to be necessary to disambiguate the potentially noisy local measurements. In our research we have worked on improving feature extraction, proposing novel blends of invariant geometric- and appearance- based features, as well as grouping algorithms that allow for the efficient construction of optimal assemblies of local features.

Regarding aspect (b) we have worked on learning scoring functions for detection with deformable models that can exploit the developed low-level representations, while also being amenable to efficient optimization. Our works in this direction build on the graph-based framework to construct models that reflect the shape properties of the structure being modeled. We have used discriminative learning to exploit boundary- and symmetry-based representations for the construction of hierarchical models for shape detection, while for medical images we have developed methods for the end-to-end discriminative training of deformable contour models that combine low-level descriptors with contour-based organ boundary representations.

Regarding aspect (c) we have developed algorithms which implement top-down/bottom-up computation both in deterministic and stochastic optimization. The main idea is that 'bottom-up', image-based guidance is necessary for efficient detection, while 'top-down', object-based knowledge can disambiguate and help reliably interpret a given image; a combination of both modes of operation is necessary to combine accuracy with efficiency. In particular we have developed novel techniques for object detection that employ combinatorial optimization tools (A* and Branch-and-Bound) to tame the combinatorial complexity, achieving a best-case performance that is logarithmic in the number of pixels.

In the long run we aim at scaling up shape-based methods to 3D detection and pose estimation and large-scale object detection. One aspect which seems central to this is the development of appropriate mid-level representations. This is a problem that has received increased interest lately in the 2D case and is relatively mature, but in 3D it has been pursued primarily through ad-hoc schemes. We anticipate that questions pertaining to part sharing in 3D will be addressed most successfully by relying on explicit 3D representations. On the one hand depth sensors, such as Microsoft's Kinect, are now cheap enough to bring surface modeling and matching into the mainstream of computer vision - so these advances may be directly exploitable at test time for detection. On the other hand, even if we do not use depth information at test time, having 3D information can simplify the modeling task during training. In on-going work with collaborators we have started exploring combinations of such aspects, namely (i) the use of surface analysis tools to match surfaces from depth sensors (ii) using branch-and-bound for efficient inference in 3D space and (iii) groupwise-registration to build statistical 3D surface models. In the coming years we intend to pursue a tighter integration of these different directions for scalable 3D object recognition.

3.2. Machine Learning & Structured Prediction

The foundation of statistical inference is to learn a function that minimizes the expected loss of a prediction with respect to some unknown distribution

$$\mathcal{R}(f) = \int \ell(f, x, y) dP(x, y), \tag{2}$$

where $\ell(f, x, y)$ is a problem specific loss function that encodes a penalty for predicting $f(x)$ when the correct prediction is y. In our case, we consider x to be a medical image, and y to be some prediction, e.g. the segmentation of a tumor, or a kinematic model of the skeleton. The loss function, ℓ, is informed by the costs associated with making a specific misprediction. As a concrete example, if the true spatial extent of a tumor is encoded in y, $f(x)$ may make mistakes in classifying healthy tissue as a tumor, and mistakes in classifying diseased tissue as healthy. The loss function should encode the potential physiological damage resulting from erroneously targeting healthy tissue for irradiation, as well as the risk from missing a portion of the tumor.

A key problem is that the distribution P is unknown, and any algorithm that is to estimate f from labeled training examples must additionally make an implicit estimate of P. A central technology of empirical inference is to approximate $\mathcal{R}(f)$ with the empirical risk,

$$\mathcal{R}(f) \approx \widehat{\mathcal{R}}(f) = \frac{1}{n} \sum_{i=1}^{n} \ell(f, x_i, y_i), \tag{3}$$

which makes an implicit assumption that the training samples (x_i, y_i) are drawn i.i.d. from P. Direct minimization of $\widehat{\mathcal{R}}(f)$ leads to overfitting when the function class $f \in \mathcal{F}$ is too rich, and regularization is required:

$$\min_{f \in \mathcal{F}} \lambda \Omega(\|f\|) + \widehat{\mathcal{R}}(f), \tag{4}$$

where Ω is a monotonically increasing function that penalizes complex functions.

Equation Eq. 4 is very well studied in classical statistics for the case that the output, $y \in \mathcal{Y}$, is a binary or scalar prediction, but this is not the case in most medical imaging prediction tasks of interest. Instead, complex interdependencies in the output space leads to difficulties in modeling inference as a binary prediction problem. One may attempt to model e.g. tumor segmentation as a series of binary predictions at each voxel in a medical image, but this violates the i.i.d. sampling assumption implicit in Equation Eq. 3. Furthermore, we typically gain performance by appropriately modeling the inter-relationships between voxel predictions, e.g. by incorporating pairwise and higher order potentials that encode prior knowledge about the problem domain. It is in this context that we develop statistical methods appropriate to structured prediction in the medical imaging setting.

3.3. Self-Paced Learning with Missing Information

Many tasks in artificial intelligence are solved by building a model whose parameters encode the prior domain knowledge and the likelihood of the observed data. In order to use such models in practice, we need to estimate its parameters automatically using training data. The most prevalent paradigm of parameter estimation is supervised learning, which requires the collection of the inputs x_i and the desired outputs y_i. However, such an approach has two main disadvantages. First, obtaining the ground-truth annotation of high-level applications, such as a tight bounding box around all the objects present in an image, is often expensive. This prohibits the use of a large training dataset, which is essential for learning the existing complex models. Second, in many applications, particularly in the field of medical image analysis, obtaining the ground-truth annotation may not be feasible. For example, even the experts may disagree on the correct segmentation of a microscopical image due to the similarities between the appearance of the foreground and background.

In order to address the deficiencies of supervised learning, researchers have started to focus on the problem of parameter estimation with data that contains hidden variables. The hidden variables model the missing information in the annotations. Obtaining such data is practically more feasible: image-level labels ('contains car','does not contain person') instead of tight bounding boxes; partial segmentation of medical images. Formally, the parameters **w** of the model are learned by minimizing the following objective:

$$\min_{\mathbf{w}\in\mathcal{W}} R(\mathbf{w}) + \sum_{i=1}^{n} \Delta(y_i, y_i(\mathbf{w}), h_i(\mathbf{w})). \tag{5}$$

Here, $\mathcal{W}$ represents the space of all parameters, n is the number of training samples, $R(\cdot)$ is a regularization function, and $\Delta(\cdot)$ is a measure of the difference between the ground-truth output y_i and the predicted output and hidden variable pair $(y_i(\mathbf{w}), h_i(\mathbf{w}))$.

Previous attempts at minimizing the above objective function treat all the training samples equally. This is in stark contrast to how a child learns: first focus on easy samples ('learn to add two natural numbers') before moving on to more complex samples ('learn to add two complex numbers'). In our work, we capture this intuition using a novel, iterative algorithm called self-paced learning (SPL). At an iteration t, SPL minimizes the following objective function:

$$\min_{\mathbf{w}\in\mathcal{W},\mathbf{v}\in\{0,1\}^n} R(\mathbf{w}) + \sum_{i=1}^{n} v_i\Delta(y_i, y_i(\mathbf{w}), h_i(\mathbf{w})) - \mu_t \sum_{i=1}^{n} v_i. \tag{6}$$

Here, samples with $v_i = 0$ are discarded during the iteration t, since the corresponding loss is multiplied by 0. The term μ_t is a threshold that governs how many samples are discarded. It is annealed at each iteration, allowing the learner to estimate the parameters using more and more samples, until all samples are used. Our results already demonstrate that SPL estimates accurate parameters for various applications such as image classification, discriminative motif finding, handwritten digit recognition and semantic segmentation. We will investigate the use of SPL to estimate the parameters of the models of medical imaging applications, such as segmentation and registration, that are being developed in the GALEN team. The ability to handle missing information is extremely important in this domain due to the similarities between foreground and background appearances (which results in ambiguities in annotations). We will also develop methods that are capable of minimizing more general loss functions that depend on the (unknown) value of the hidden variables, that is,

$$\min_{\mathbf{w}\in\mathcal{W},\theta\in\Theta} R(\mathbf{w}) + \sum_{i=1}^{n} \sum_{h_i\in\mathcal{H}} \Pr(h_i|x_i, y_i; \theta)\Delta(y_i, h_i, y_i(\mathbf{w}), h_i(\mathbf{w})). \tag{7}$$

Here, θ is the parameter vector of the distribution of the hidden variables h_i given the input x_i and output y_i, and needs to be estimated together with the model parameters **w**. The use of a more general loss function will allow us to better exploit the freely available data with missing information. For example, consider the case where y_i is a binary indicator for the presence of a type of cell in a microscopical image, and h_i is a tight bounding box around the cell. While the loss function $\Delta(y_i, y_i(\mathbf{w}), h_i(\mathbf{w}))$ can be used to learn to classify an image as containing a particular cell or not, the more general loss function $\Delta(y_i, h_i, y_i(\mathbf{w}), h_i(\mathbf{w}))$ can be used to learn to detect the cell as well (since h_i models its location)

3.4. Discrete Biomedical Image Perception

A wide variety of tasks in medical image analysis can be formulated as discrete labeling problems. In very simple terms, a discrete optimization problem can be stated as follows: we are given a discrete set of variables $\mathcal{V}$, all of which are vertices in a graph $\mathcal{G}$. The edges of this graph (denoted by $\mathcal{E}$) encode the variables' relationships. We are also given as input a discrete set of labels $\mathcal{L}$. We must then assign one label from $\mathcal{L}$ to each variable in $\mathcal{V}$. However, each time we choose to assign a label, say, x_{p_1} to a variable p_1, we are forced to pay a price according to the so-called *singleton* potential function $g_p(x_p)$, while each time we choose to assign a pair of labels, say, x_{p_1} and x_{p_2} to two interrelated variables p_1 and p_2 (two nodes that are connected by an edge in the graph $\mathcal{G}$), we are also forced to pay another price, which is now determined by the so called *pairwise* potential function $f_{p_1 p_2}(x_{p_1}, x_{p_2})$. Both the singleton and pairwise potential functions are problem specific and are thus assumed to be provided as input.

Our goal is then to choose a labeling which will allow us to pay the smallest total price. In other words, based on what we have mentioned above, we want to choose a labeling that minimizes the sum of all the MRF potentials, or equivalently the MRF energy. This amounts to solving the following optimization problem:

$$\arg\min_{\{x_p\}} \mathcal{P}(g, f) = \sum_{p \in \mathcal{V}} g_p(x_p) + \sum_{(p_1, p_2) \in \mathcal{E}} f_{p_1 p_2}(x_{p_1}, x_{p_2}). \tag{8}$$

The use of such a model can describe a number of challenging problems in medical image analysis. However these simplistic models can only account for simple interactions between variables, a rather constrained scenario for high-level medical imaging perception tasks. One can augment the expression power of this model through higher order interactions between variables, or a number of cliques $\{C_i, i \in [1, n] = \{\{p_{i^1}, \cdots, p_{i^{|C_i|}}\}\}$ of order $|C_i|$ that will augment the definition of $\mathcal{V}$ and will introduce hyper-vertices:

$$\arg\min_{\{x_p\}} \mathcal{P}(g, f) = \sum_{p \in \mathcal{V}} g_p(x_p) + \sum_{(p_1, p_2) \in \mathcal{E}} f_{p_1 p_2}(x_{p_1}, x_{p_2}) + \sum_{C_i \in \mathcal{E}} f_{p_1 \cdots p_n}(x_{p_{i^1}}, \cdots, p_{x_{i^{|C_i|}}}). \tag{9}$$

where $f_{p_1 \cdots p_n}$ is the price to pay for associating the labels $(x_{p_{i^1}}, \cdots, p_{x_{i^{|C_i|}}})$ to the nodes $(p_1 \cdots p_{i^{|C_i|}})$. Parameter inference, addressed by minimizing the problem above, is the most critical aspect in computational medicine and efficient optimization algorithms are to be evaluated both in terms of computational complexity as well as of inference performance. State of the art methods include deterministic and non-deterministic annealing, genetic algorithms, max-flow/min-cut techniques and relaxation. These methods offer certain strengths while exhibiting certain limitations, mostly related to the amount of interactions which can be tolerated among neighborhood nodes. In the area of medical imaging where domain knowledge is quite strong, one would expect that such interactions should be enforced at the largest scale possible.

4. Application Domains

4.1. Representation Learning for Network Biology

Participants: Fragkiskos Malliaros, Abdulkadir Çelikkanat (Collaboration: Duong Nguyen, UC San Diego)

Networks (or graphs) are ubiquitous in the domain of biology, as many biological systems can naturally be mapped to graph structures. Characteristic examples include protein-protein interaction and gene regulatory networks. To this extend, machine learning on graphs is an important task with many practical applications in network biology. For example, in the case on protein-protein interaction networks, predicting the function of a protein is a key task that assigns biochemical roles to proteins. The main challenge here is to find appropriate representations of the graph structure, in order to be easily exploited by machine learning models. The traditional approach to the problem was relying on the extraction of “hand-crafted" discriminating features that encode information about the graph, based on user-defined heuristics. Nevertheless, this approach has demonstrated severe limitations, as the learning process heavily depends on the manually extracted features. To this end, feature (or representation) learning techniques can be used to automatically learn to encode the graph structure into low-dimensional feature vectors – which can later be used in learning tasks. Our goal here is to develop a systematic framework for large-scale representation learning on biological graphs. Our approach takes advantage of the clustering structure of these networks, to further enhance the ability of the learned features to capture intrinsic structural properties.

4.2. Breast tomosynthesis

Participants: Emilie Chouzenoux, Jean-Christophe Pesquet, Maissa Sghaier (collaboration G. Palma, GE Healthcare)

Breast cancer is the most frequently diagnosed cancer for women. Mammography is the most used imagery tool for detecting and diagnosing this type of cancer. Since it consists of a 2D projection method, this technique is sensitive to geometrical limitations such as the superimposition of tissues which may reduce the visibility of lesions or make even appear false structures which are interpreted by radiologists as suspicious signs. Digital breast tomosynthesis allows these limitations to be circumvented. This technique is grounded on the acquisition of a set of projections with a limited angle view. Then, a 3D estimation of the sensed object is performed from this set of projections, so reducing the overlap of structures and improving the visibility and detectability of lesions possibly present in the breast. The objective of our work is to develop a high quality reconstruction methodology where the full pipeline of data processing will be modeled.

4.3. Inference of gene regulatory networks

Participants: Jean-Christophe Pesquet (collaboration A. Pirayre and L. Duval, IFPEN)

The discovery of novel gene regulatory processes improves the understanding of cell phenotypic responses to external stimuli for many biological applications, such as medicine, environment or biotechnologies. To this purpose, transcriptomic data are generated and analyzed from DNA microarrays or more recently RNAseq experiments. They consist in genetic expression level sequences obtained for all genes of a studied organism placed in dierent living conditions. From these data, gene regulation mechanisms can be recovered by revealing topological links encoded in graphs. In regulatory graphs, nodes correspond to genes. A link between two nodes is identified if a regulation relationship exists between the two corresponding genes. In our work, we propose to address this network inference problem with recently developed techniques pertaining to graph optimization. Given all the pairwise gene regulation information available, we propose to determine the presence of edges in the considered GRN by adopting an energy optimization formulation integrating additional constraints. Either biological (information about gene interactions) or structural (information about node connectivity) a priori are considered to restrict the space of possible solutions. Different priors lead to different properties of the global cost function, for which various optimization strategies, either discrete and continuous, can be applied.

4.4. Imaging biomarkers and characterization for chronic lung diseases

Participants: Guillaume Chassagnon, Maria Vakalopoulou (in collaboration with Evangelia Zacharaki and Nikos Paragios: University of Patras; Therapanacea)

Diagnosis and staging of chronic lung diseases is a major challenge for both patient care and approval of new treatments. Among imaging techniques, computed tomography (CT) is the gold standard for in vivo morphological assessment of lung parenchyma currently offering the highest spatial resolution in chronic lung diseases. Although CT is widely used its optimal use in clinical practice and as an endpoint in clinical trials remains controversial. Our goal is to develop quantitative imaging biomarkers that allow (i)severity assessment (based on the correlation to functional and clinical data) and (ii) monitoring the disease progression. In the current analysis we focus on scleroderma and cystic fibrosis as models for restrictive and obstructive lung disease, respectively. Two different approaches are investigated: disease assessment by deep convolutional neural networks and assessment of the regional lung elasticity through deformable registration. This work is in collaboration with the Department of Radiology, Cochin Hospital, Paris.

4.5. Imaging radiomics and genes to assess immunotherapy

Participants: Roger Sun, Théo Estienne, Enzo Batistella, Maria Vakalopoulou (in collaboration with Éric Deutsch and Nikos Paragios: Institut de Cancérologie Gustave Roussy, Therapanacea)
Because responses of patients with cancer to immunotherapy can vary in success, innovative predictors of response to treatment are urgently needed to improve treatment outcomes. We aimed to develop and independently validate a radiomics-based biomarker of tumour-infiltrating CD8 cells in patients included in phase 1 trials of anti-programmed cell death protein (PD)-1 or anti-programmed cell death ligand 1 (PD-L1) monotherapy. We also aimed to evaluate the association between the biomarker, and tumour immune phenotype and clinical outcomes of these patients. This work is in collaboration with the Institut de Cancérologie Gustave Roussy Paris.

4.6. Restoration of old video archives

Participants: Emilie Chouzenoux, Jean-Christophe Pesquet (collaboration F. Abboud, WITBE, J.-H. Chenot and L. Laborelli, INA)
The last century has witnessed an explosion in the amount of video data stored with holders such as the National Audiovisual Institute whose mission is to preserve and promote the content of French broadcast programs. the cultural impact of these records, their value is increased due to commercial reexploitation through recent visual media. However, the perceived quality of the old data fails to satisfy the current public demand. The purpose of our work is to propose new methods for restoring video sequences supplied from television archive documents, using modern optimization techniques with proven convergence properties [21], [50], [3].

4.7. Development of a heart ventricle vessel generation model for perfusion analysis

Participant: Hugues Talbot (collaboration with L. Najman ESIEE Paris, I. Vignon-Clementel, REO Team, Inria, Leo Grady, Heartflow Inc.)
Cardio-vascular diseases are the leading cause of mortality in the world. Understanding these diseases is still a current, challenging and essential research project. The leading cause of heart malfunction are stenoses causing ischemia in the coronary vessels. Current CT and MRI technology allow to assess coronary diseases but are typically invasive, requiring catheterization and relatively toxic contrast agents injection. With Heartflow, a US based company, we have in the past worked to use image-based exams only, limiting the use of contrast agents and in many cases eliminating catheterisation. Heartflow is current the market leader in non-invasive coronary exams.

Unfortunately, current imaging technology is unable to assess the full length of coronary vessels. CT is limited to a resolution of about 1mm, whereas coronary vessels can be much smaller. Blood perfusion throughout the heart muscle can provide insight regarding coronary health in areas that CT or MRI cannot assess. Perfusion imaging with PET or a Gamma camera, the current gold standard, is an invasive technology requiring the use of radioactive tracers.

With heartflow and with the REO team of Inria, we have investigated patient-specific vessel generation models together with porous model simulations in order to propose a forward model of perfusion imaging, based on the known patient data, computer flow dynamic simulations as well as experimental data consistent with known vessel and heart muscle physiology. The objective of this work is to solve the inverse problem of locating and assessing coronary diseases even though the affected vessels are too small to be imaged directly.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

The work on graph-based text categorization by F. Malliaros et al. [39] has received the best paper award at the 12th NAACL-HLT Workshop on Graph-Based Natural Language Processing (TextGraphs), held in New Orleans, Louisiana in June 2018.

Riza Alp Güler obtained the 2nd place at Prix du Doctorant for the Doctoral School STIC of Univ. Paris Saclay.

M. Papadomanolaki and M. Vakalopoulou got the 2nd place at the Earth Observation Challenge organised by Digital Globe and ESA for the project UrbanMonitor: Mapping Changes in Urban Environments towards Resilient Cities and Urban Sustainability. http://blog.digitalglobe.com/news/earth-observation-challenge-the-three-winners/

A. Pirayre whose PhD thesis was advised by J.-C. Pesquet received the Yves Chauvin PhD award (IFPEN).

Our M.Sc. program in Data Sciences and Business Analytics (with ESSEC Business School) was ranked 4th worldwide in the QS World University Rankings.

6. New Software and Platforms

6.1. Platforms

6.1.1. The Proximity Operator Repository

Participants: Emilie Chouzenoux and Jean-Christophe Pesquet (in collaboration with Giovanni Chierchia, Univ. Paris Est, and Patrick Combettes, North Carolina State University).
link: http://proximity-operator.net/
Proximity operators have become increasingly important tools as basic building blocks of proximal splitting algorithms, a class of algorithms that decompose complex composite convex optimization methods into simple steps involving one of the functions present in the model. This website provides formulas for efficiently computing the proximity operator of various functions, along with the associated codes.

6.1.2. Kymatio

Participant: Edouard Oyallon.
link: http://www.kymat.io
The website shares the software `Kymatio` for Scattering Transform, that computes cascade of wavelets and modulus non-linearity. The codes have been optimized for GPUs and work on the open-source framework *PyTorch*.

6.1.3. Topical Node Embeddings

Participant: Abdulkadir Çelikkanat
link: https://abdcelikkanat.github.io/projects/TNE/

TNE learns node representations on graphs combining node and topic information, as described in [45]. The framework has been implemented in Python and has been built upon widely used modules, including `networkx`, `scipy`, `gensim` and `scikit-learn`.

6.1.4. BiasedWalk: Learning latent node features with biased random walks

Participants: Fragkiskos Malliaros and Duong Nguyen (UC San Diego)
link: https://github.com/duong18/BiasedWalk/
The BiasedWalk network representation learning algorithm, computes latent node features on graphs based on biased random walks [33]. The framework has been implemented in Python and has been built upon widely used modules, including `networkx`, `numpy`, `gensim` and `scikit-learn`.

6.1.5. DiffuGreedy: Influence maximization in complex networks based on diffusion cascades

Participants: Fragkiskos Malliaros, Georgios Panagopoulos and Michalis Vazirgiannis (École Polytechnique)
link: https://goo.gl/GpfCVZ
The DiffuGreedy is an algorithm for influence maximization in complex networks, that is based on diffusion cascades. The algorithm implements the methodology described in [35]. It has been implemented in Python and has been built upon widely used modules, including `networkx`, `igraph`, `numpy` and `pandas`.

6.1.6. Graph-based text categorization

Participants: Fragkiskos Malliaros, Konstantinos Skianis and Michalis Vazirgiannis (École Polytechnique)
link: https://github.com/y3nk0/Graph-Based-TC/
Graph-based TC is a framework for text categorization that relies on a graph representation of documents. The framework uses various graph centrality criteria to determine the importance of a term within a document. It also makes use of word embeddings to further boost the performance of graph-based methods. The algorithm implements the methodology described in [39]. It has been implemented in Python and has been built upon widely used modules, including `networkx`, `igraph`, `numpy` and `scikit-learn`.

6.1.7. The PINK image library

Participant: Hugues Talbot
link: http://ibipio.hu/joomla/
The PINK image library is a general-purpose, open-source, portable image processing library specializing in discrete geometry and mathematical morphology. It is the result of several decades of research in these domains and features state-of-the art algorithmic implementation of both classical and leading edge DG and MM operators. These include nD parallel thinning and skeletonization methods and efficient hierarchical morphological transforms.

6.1.8. The Vivabrain AngioTK toolkit

Participant: Hugues Talbot
link: https://github.com/vivabrain/angiotk
AngioTK is a toolkit supported by Kitware (the authors of VTK) for the filtering, segmentation, generation and simulation of blood vessels. It was started in the context of the Vivabrain ANR project in 2012, but continues with the same as well as new partners. Applications are numerous, from the simulation and understanding of perfusion (see associated theme) to the simulation of realistic blood flow MRI images with associated ground truth, via the generation of blood vessel atlases.

7. New Results

7.1. Invertible Deep Networks

Participant: Edouard Oyallon (in collaboration with J.H. Jacobsen and A. Smeulders, Instituut voor Informatica)

It is widely believed that the success of deep convolutional networks is based on progressively discarding uninformative variability about the input with respect to the problem at hand. This is supported empirically by the difficulty of recovering images from their hidden representations, in most commonly used network architectures. In this paper we show via a one-to-one mapping that this loss of information is not a necessary condition to learn representations that generalize well on complicated problems, such as ImageNet. Via a cascade of homeomorphic layers, we build the i-RevNet, a network that can be fully inverted up to the final projection onto the classes, i.e. no information is discarded. Building an invertible architecture is difficult, for one, because the local inversion is ill-conditioned, we overcome this by providing an explicit inverse. An analysis of i-RevNet's learned representations suggests an alternative explanation for the success of deep networks by a progressive contraction and linear separation with depth. To shed light on the nature of the model learned by the i-RevNet we reconstruct linear interpolations between natural image representations [28].

7.2. Compression of CNNs inputs

Participant: Edouard Oyallon (in collaboration with E. Belilovsky, DIRO, Montréal, S. Zagoruyko, WILLOW, Inria Paris and M. Valko, SEQUEL, Inria Lille)

Typical inputs of CNNs are highly redundant and could be potentially reduced. We study the first-order scattering transform as a candidate for reducing the signal processed by a convolutional neural network (CNN). We study this transformation and show theoretical and empirical evidence that in the case of natural images and sufficiently small translation invariance, this transform preserves most of the signal information needed for classification while substantially reducing the spatial resolution and total signal size. We show that cascading a CNN with this representation performs on par with ImageNet classification models commonly used in downstream tasks such as the ResNet-50. We subsequently apply our trained hybrid ImageNet model as a base model on a detection system, which has typically larger image inputs. On Pascal VOC and COCO detection tasks we deliver substantial improvements in the inference speed and training memory consumption compared to models trained directly on the input image [34].

7.3. Interstitial lung disease segmentation

Participants: Guillaume Chassagnon, Norbert Bus, Rafael Marini Silva, Evangelia Zacharaki, Maria Vakalopoulou (in collaboration with Marie-Pierre Revel and Nikos Paragios: AP-HP - Hopital Cochin Broca Hotel Dieu; Therapanacea)

Interstitial lung diseases (ILD) encompass a large spectrum of diseases sharing similarities in their physiopathology and computed tomography (CT) appearance. In the work [42], we propose the adaption of a deep convolutional encoder-decoder (CED) that has shown high accuracy for image segmentation. Such architectures require annotation of the total region with pathological findings. This is difficult to acquire, due to uncertainty in the definition and extent of disease patterns and the need of significant human effort, especially for large datasets. Therefore, often current methods use patch-based implementations of convolutional neural networks, which however tend to produce spatially inhomogeneous segmentations due to their local contextual view. We exploit the advantages of both architectures by using the output of a patch-based classifier as a prior to a CED.

Moreover, in order to deal with the limited available datasets that are available, in [41], we introduce a novel multi-network architecture that exploits domain knowledge to address those challenges. The proposed architecture consists of multiple deep neural networks that are trained after co-aligning multiple anatomies through multi-metric deformable registration. This multi-network architecture can be trained with fewer examples and leads to better performance, robustness and generalization through consensus. Comparable to human accuracy, highly promising results on the challenging task of interstitial lung disease segmentation demonstrate the potential of our approach.

7.4. Image Registration with 3D Convolutional Neural Networks

Participants: Stergios Christodoulidis, Mihir Sahasrabudhe, Guillaume Chassagnon, Maria Vakalopoulou (in collaboration with Stavroula Mougiakakou and Marie-Pierre Revel and Nikos Paragios: University of Bern; AP-HP - Hopital Cochin Broca Hotel Dieu; Therapanacea)

Image registration and in particular deformable registration methods are pillars of medical imaging. Inspired by the recent advances in deep learning, we propose in this paper, [25] we proposed a new deep learning based and unsupervised method for image registration. In partucular, a novel convolutional neural network architecture that couples linear and deformable registration within a unified architecture endowed with near real-time performance. Our framework is modular with respect to the global transformation component, as well as with respect to the similarity function while it guarantees smooth displacement fields. We evaluate the performance of our network on the challenging problem of MRI lung registration, and demonstrate superior performance with respect to state of the art elastic registration methods. The proposed deformation (between inspiration & expiration) was considered within a clinically relevant task of interstitial lung disease (ILD) classification and showed promising results.

7.5. Radiomics for response to immunotherapy

Participants: Roger Sun, Maria Vakalopoulou (in collaboration with Elaine Johanna Limkin, Laurent Dercle, Stéphane Champiat, Shan Rong Han, Loic Verlingue, David Brandao, Andrea Lancia, Samy Ammari, Antoine Hollebecque, Jean-Yves Scoazec, Aurélien Marabelle, Christophe Massard, Jean-Charles Soria, Charlotte Robert, Nikos Paragios, Eric Deutsch, Charles Ferté: Institute Gustave Roussy; Therapanacea)

Because responses of patients with cancer to immunotherapy can vary in success, innovative predictors of response to treatment are urgently needed to improve treatment outcomes. In this retrospective multicohort work [17], we used four independent cohorts of patients with advanced solid tumours to develop and validate a radiomic signature predictive of immunotherapy response by combining contrast-enhanced CT images and RNA-seq genomic data from tumour biopsies to assess CD8 cell tumour infiltration. To develop the radiomic signature of CD8 cells, we used the CT images and RNA sequencing data of 135 patients with advanced solid malignant tumours who had been enrolled into the MOSCATO trial between May 1, 2012, and March 31, 2016, in France (training set). The genomic data, which are based on the CD8B gene, were used to estimate the abundance of CD8 cells in the samples and data were then aligned with the images to generate the radiomic signatures. The concordance of the radiomic signature (primary endpoint) was validated in a Cancer Genome Atlas [TGCA] database dataset including 119 patients who had available baseline preoperative imaging data and corresponding transcriptomic data on June 30, 2017. From 84 input variables used for the machine-learning method (78 radiomic features, five location variables, and one technical variable), a radiomics-based predictor of the CD8 cell expression signature was built by use of machine learning (elastic-net regularised regression method). Two other independent cohorts of patients with advanced solid tumours were used to evaluate this predictor. The immune phenotype internal cohort (n=100), were randomly selected from the Gustave Roussy Cancer Campus database of patient medical records based on previously described, extreme tumour-immune phenotypes: immune-inflamed (with dense CD8 cell infiltration) or immune-desert (with low CD8 cell infiltration), irrespective of treatment delivered; these data were used to analyse the correlation of the immune phenotype with this biomarker. Finally, the immunotherapy-treated dataset (n=137) of patients recruited from Dec 1, 2011, to Jan 31, 2014, at the Gustave Roussy Cancer Campus, who had been treated with anti-PD-1 and anti-PD-L1 monotherapy in phase 1 trials, was used to assess the predictive value of this biomarker in terms of clinical outcome.

7.6. Semantic Segmentation Techniques for Brain Tumor Patients

Participants: Siddhartha Chandra, Théo Estienne, Roger Sun, Enzo Battistella, Maria Vakalopoulou (in collaboration with Charlotte Robert, Nikos Paragios, Eric Deutsch: Institute Gustave Roussy; Therapanacea)

In this work [23] we propose a novel deep learning based pipeline for the task of brain tumor segmentation. Our pipeline consists of three primary components: (i) a preprocessing stage that exploits histogram standardization to mitigate inaccuracies in measured brain modalities, (ii) a first prediction stage that uses the V-Net deep learning architecture to output dense, per voxel class probabilities, and (iii) a prediction refinement stage that uses a Conditional Random Field (CRF) with a bilateral filtering objective for better context awareness. Additionally, we compare the V-Net architecture with a custom 3D Residual Network architecture, trained on a multi-view strategy, and our ablation experiments indicate that V-Net outperforms the 3D ResNet-18 with all bells and whistles, while fully connected CRFs as post processing, boost the performance of both networks. We report competitive results on the BraTS 2018 validation and test set as also summarized on [52].

7.7. Demystification of AI-driven medical image interpretation

Participants: Maria Vakalopoulou (in collaboration with P. Savadjiev, J. Chong, A. Dohan, C. Reinhold, B. Gallix: McGill University; Therapanacea)

The recent explosion of 'big data' has ushered in a new era of artificial intelligence (AI) algorithms in every sphere of technological activity, including medicine, and in particular radiology. However, the recent success of AI in certain flagship applications has, to some extent, masked decades-long advances in computational technology development for medical image analysis. In this work [16], we provide an overview of the history of AI methods for radiological image analysis in order to provide a context for the latest developments. We review the functioning, strengths and limitations of more classical methods as well as of the more recent deep learning techniques. We discuss the unique characteristics of medical data and medical science that set medicine apart from other technological domains in order to highlight not only the potential of AI in radiology but also the very real and often overlooked constraints that may limit the applicability of certain AI methods. Finally, we provide a comprehensive perspective on the potential impact of AI on radiology and on how to evaluate it not only from a technical point of view but also from a clinical one, so that patients can ultimately benefit from it.

7.8. Semantic Segmentation for Remote Sensing Data

Participants: Maria Papadomanolaki, Maria Vakalopoulou (in collaboration with Christina Karakizi, Georgia Antoniou, Konstantinos Karantzalos, Nikos Paragios; National Technical University of Athens, Therapanacea)

Detailed, accurate and frequent land cover mapping is a prerequisite for several important geospatial applications and the fulfilment of current sustainable development goals. This work [9] introduces a methodology for the classification of annual high-resolution satellite data into several detailed land cover classes. In particular, a nomenclature with 27 different classes was introduced based on CORINE Land Cover (CLC) Level-3 categories and further analysing various crop types. Without employing cloud masks and/or interpolation procedures, we formed experimental datasets of Landsat-8 (L8) images with gradually increased cloud cover in order to assess the influence of cloud presence on the reference data and the resulting classification accuracy. The performance of shallow kernel-based and deep patch-based machine learning classification frameworks was evaluated. Quantitatively, the resulting overall accuracy rates differed within a range of less than 3%; however, maps produced based on Support Vector Machines (SVM) were more accurate across class boundaries and the respective framework was less computationally expensive compared to the applied patch-based deep Convolutional Neural Network (CNN). Further experimental results and analysis indicated that employing all multitemporal images with up to 30% cloud cover delivered relatively higher overall accuracy rates as well as the highest per-class accuracy rates. Moreover, by selecting 70% of the top-ranked features after applying a feature selection strategy, slightly higher accuracy rates were achieved. A detailed discussion of the quantitative and qualitative evaluation outcomes further elaborates on the performance of all considered classes and highlights different aspects of their spectral behaviour and separability.

Moreover, semantic segmentation is a mainstream method in several remote sensing applications based on very-high-resolution data, achieving recently remarkable performance by the use of deep learning and more specifically, pixel-wise dense classification models. In this work [36], we exploit the use of a relatively deep architecture based on repetitive downscale upscale processes that had been previously employed for human pose estimation. By integrating such a model, we are aiming to capture low-level details, such as small objects, object boundaries and edges. Experimental results and quantitative evaluation has been performed on the publicly available ISPRS (WGIII/4) benchmark dataset indicating the potential of the proposed approach.

7.9. BRANE Clust: Cluster-Assisted Gene Regulatory Network Inference Refinement

Participants: Jean-Christophe Pesquet (in collaboration with Aurélie Pirayre, IFP Energies nouvelles, Camille Couprie, Facebook Research, Laurent Duval, IFP Energies nouvelles)

Discovering meaningful gene interactions is crucial for the identification of novel regulatory processes in cells. Building accurately the related graphs remains challenging due to the large number of possible solutions from available data. Nonetheless, enforcing a priori on the graph structure, such as modularity, may reduce network indeterminacy issues. BRANE Clust (Biologically-Related A priori Network Enhancement with Clustering) refines gene regulatory network (GRN) inference thanks to cluster information. It works as a post-processing tool for inference methods (i.e. CLR, GENIE3). In BRANE Clust, the clustering is based on the inversion of a system of linear equations involving a graph-Laplacian matrix promoting a modular structure. Our approach [14] is validated on DREAM4 and DREAM5 datasets with objective measures, showing significant comparative improvements. We provide additional insights on the discovery of novel regulatory or co-expressed links in the inferred Escherichia coli network evaluated using the STRING database. The comparative pertinence of clustering is discussed computationally (SIMoNe, WGCNA, X-means) and biologically (RegulonDB).

7.10. Proximity Operators of Discrete Information Divergences

Participants: Jean-Christophe Pesquet (in collaboration with Mireille El Gheche, EPFL, Giovanni Chierchia, ESIEE Paris)

Information divergences allow one to assess how close two distributions are from each other. Among the large panel of available measures, a special attention has been paid to convex ϕ-divergences, such as Kullback-Leibler, Jeffreys-Kullback, Hellinger, Chi-Square, Renyi, and I_α divergences. While ϕ-divergences have been extensively studied in convex analysis, their use in optimization problems often remains challenging. In this regard, one of the main shortcomings of existing methods is that the minimization of ϕ-divergences is usually performed with respect to one of their arguments, possibly within alternating optimization techniques. In this paper, we overcome this limitation by deriving new closed-form expressions for the proximity operator of such two-variable functions. This makes it possible to employ standard proximal methods for efficiently solving a wide range of convex optimization problems involving ϕ-divergences. In addition, we show that these proximity operators are useful to compute the epigraphical projection of several functions of practical interest. The proposed proximal tools are numerically validated in the context of optimal query execution within database management systems, where the problem of selectivity estimation plays a central role. Experiments are carried out on small to large scale scenarios [6].

7.11. Stochastic quasi-Fejèr block-coordinate fixed point iterations with random sweeping

Participants: Jean-Christophe Pesquet (in collaboration with Patrick Combettes, North Caroline State University)

Our previous work investigated the almost sure weak convergence of block-coordinate fixed point algorithms and discussed their applications to nonlinear analysis and optimization. This algorithmic framework features random sweeping rules to select arbitrarily the blocks of variables that are activated over the course of the iterations and it allows for stochastic errors in the evaluation of the operators. The present paper establishes results on the mean-square and linear convergence of the iterates. Applications to monotone operator splitting and proximal optimization algorithms are presented.

7.12. Rational optimization for nonlinear reconstruction with approximate ℓ_0 penalization

Participants: Marc Castella, Arthur Marmin, Jean-Christophe Pesquet

Recovering nonlinearly degraded signal in the presence of noise is a challenging problem. In this work, this problem is tackled by minimizing the sum of a non convex least-squares fit criterion and a penalty term. We assume that the nonlinearity of the model can be accounted for by a rational function. In addition, we suppose that the signal to be sought is sparse and a rational approximation of the ℓ_0 pseudo-norm thus constitutes a suitable penalization. The resulting composite cost function belongs to the broad class of semi-algebraic functions. To find a globally optimal solution to such an optimization problem, it can be transformed into a generalized moment problem, for which a hierarchy of semidefinite programming relaxations can be built. Global optimality comes at the expense of an increased dimension and, to overcome computational limitations concerning the number of involved variables, the structure of the problem has to be carefully addressed. A situation of practical interest is when the nonlinear model consists of a convolutive transform followed by a componentwise nonlinear rational saturation. We then propose to use a sparse relaxation able to deal with up to several hundreds of optimized variables. In contrast with the naive approach consisting of linearizing the model, our experiments show that the proposed approach offers good performance [53].

7.13. Representation Learning on Real-World Graphs

Participants: Fragkiskos Malliaros, Abdulkadir Çelikkanat (in collaboration with Duong Nguyen, UC San Diego)

Network representation learning (NRL) methods aim to map each vertex into a low dimensional space by preserving both local and global structure of a given network. In recent years, various approaches based on random walks have been proposed to learn node embeddings – thanks to their success in several challenging problems. In this work, we have introduced two methodologies to compute latent representations of nodes based on random walks.

In particular, we have proposed BiasedWalk, an unsupervised Skip-gram-based network embedding algorithm which can preserve higher-order proximity information, as well as capture both the homophily and role equivalence relationships between nodes [33]. BiasedWalk relies on a novel node sampling procedure based on biased random walks, that can behave as actual depth-first-search and breath-first-search explorations – thus, forcing the sampling scheme to capture both role equivalence and homophily relations between nodes. Furthermore, BiasedWalk is scalable on large scale graphs, and is able to handle different types of networks structures, including (un)weighted and (un)directed ones.

Furthermore, we have introduced TNE (Topical Node Embeddings), a general framework to enhance node embeddings acquired by means of the random walk-based approaches [45]. Similar to the notion of *topical word embeddings* in the domain of Natural Language Processing, the proposed framework assigns each vertex to a topic with the favor of various statistical models and community detection methods, and then generates enhanced community representations.

We have evaluated our methods on two downstream tasks: node classification and link prediction in social, information and biological networks. The experimental results demonstrate that the biased random walks as well as the incorporation of vertex and topic embeddings outperform widely-known baseline NRL methods.

7.14. Anonymity on Directed Networks

Participants: Fragkiskos Malliaros (in collaboration with Jordi Casas-Roma and Julián Salas, Universitat Oberta de Catalunya; Michalis Vazirgiannis, École Polytechnique)

In recent years, a huge amount of social and human interaction networks have been made publicly available. Embedded within this data, there is user's private information that must be preserved before releasing the data to third parties and researchers. In this work, we have considered the problem of anonymization on directed networks. Although there are several anonymization methods for networks, most of them have explicitly been designed to work with undirected networks and they can not be straightforwardly applied when they are directed. Moreover, ignoring the direction of the edges causes important information loss on the anonymized networks in the best case. In the worst case, the direction of the edges may be used for reidentification, if it is not considered in the anonymization process. Here, we have proposed two different models for k-degree anonymity on directed networks, and we also present algorithms to fulfill these k-degree anonymity models [4]. Given a network G, we construct a k-degree anonymous network by the minimum number of edge additions. Our algorithms use multivariate micro-aggregation to anonymize the degree sequence, and then they modify the graph structure to meet the k-degree anonymous sequence. We apply our algorithms to several real datasets and demonstrate their efficiency and practical utility.

7.15. Influence Maximization in Complex Networks

Participants: Fragkiskos Malliaros (in collaboration with Michalis Vazirgiannis, George Panagopoulos, Maria-Evgenia Rossi, Bowen Shi, Christos Giatsidis, École Polytechnique; Nikolaos Tziortziotis, Université Paris-Sud)

Influence maximization in complex networks has attracted a lot of attention due to its numerous applications, including diffusion of social movements, the spread of news, viral marketing and outbreak of diseases. The objective is to discover a group of users that are able to maximize the spread of influence across a network. The seminal *greedy* algorithm developed by Kempe, Kleinberg and Tardos progressively adds new nodes to the seed set, maximizing the expected influence spread; the algorithm gives a solution to the influence maximization problem while having a good approximation ratio.

Nevertheless, one of the bottlenecks of the greedy algorithm is that it does not scale well on large scale datasets. In our work, we have proposed Matrix Influence (MATI), an efficient algorithm that can be used under both the Linear Threshold and Independent Cascade diffusion models [15]. MATI is based on the precalculation of the influence by taking advantage of the simple paths in the node's neighborhood. An extensive empirical analysis has been performed on multiple real-world datasets showing that MATI has competitive performance when compared to other well-known algorithms with regards to running time and expected influence spread.

Furthermore, the previously described greedy algorithm focuses solely on static networks. However, with the emergence of several complementary data, such as the network's temporal changes and the diffusion cascades taking place over it, novel methods have been proposed with promising results. In our work, we have introduced a simple yet effective algorithm (called DiffuGreedy) that combines the algorithmic methodology of the greedy approach with diffusion cascades [35]. We have compared it with four different prevalent influence maximization approaches, on a large scale Chinese microblogging dataset. More specifically, for comparison, we have employed methods that derive the seed set using the static network, the temporal network, the diffusion cascades, and their combination. A set of diffusion cascades from the latter part of the dataset is set aside for evaluation. The experimental evaluation has shown that the proposed DiffuGreedy outperforms widely used baseline methods in both quality of the seed set and computational efficiency.

7.16. Graph-based Text Analytics

Participants: Fragkiskos Malliaros (in collaboration with Konstantinos Skianis and Michalis Vazirgiannis, École Polytechnique)

Text categorization is a core task in a plethora of text mining applications. In our work, contrary to the traditional *Bag-of-Words* approach, we have considered the *Graph-of-Words* model in which each document is represented by a graph that encodes relationships between the different terms. Based on this formulation, we treat the term weighting task as a node ranking problem; the importance of a term is determined by the importance of the corresponding node in the graph, using node centrality criteria. We have also introduced novel graph-based weighting schemes by enriching graphs with word-embedding distances, in order to reward or penalize the importance of semantically close terms [39]. Our methods produce more discriminative feature weights for text categorization, outperforming existing frequency-based criteria – highlighting also the importance of graph-based methods in text analytics and natural language processing in general.

7.17. Auxiliary Variable Method for MCMC Algorithms in High Dimension

Participants: Emilie Chouzenoux, Jean-Christophe Pesquet (Collaboration: Yosra Marnissi, SAFRAN TECH, and Amel Benazza-Benyahia, SUP'COM Tunis)

In this work, we are interested in Bayesian inverse problems where either the data fidelity term or the prior distribution is Gaussian or driven from a hierarchical Gaussian model. Generally, Markov chain Monte Carlo (MCMC) algorithms allow us to generate sets of samples that are employed to infer some relevant parameters of the underlying distributions. However, when the parameter space is high-dimensional, the performance of stochastic sampling algorithms is very sensitive to existing dependencies between parameters. In particular, this problem arises when one aims to sample from a high-dimensional Gaussian distribution whose covariance matrix does not present a simple structure. Another challenge is the design of Metropolis–Hastings proposals that make use of information about the local geometry of the target density in order to speed up the convergence and improve mixing properties in the parameter space, while not being too computationally expensive. These two contexts are mainly related to the presence of two heterogeneous sources of dependencies stemming either from the prior or the likelihood in the sense that the related covariance matrices cannot be diagonalized in the same basis. In this work, we address these two issues. Our contribution consists of adding auxiliary variables to the model in order to dissociate the two sources of dependencies. In the new augmented space, only one source of correlation remains directly related to the target parameters, the other sources of correlations being captured by the auxiliary variables. Experimental results conducted on two practical image restoration problems indicate that adding the proposed auxiliary variables makes the sampling problem simpler, and thus the computational cost of each iteration of the Gibbs sampler is significantly reduced while ensuring good mixing properties [11].

7.18. Generation of patient-specific cardiac vascular networks

Participant: Hugues Talbot (in collaboration with C. Jaquet, L. Najman, ESIEE Paris; L. Grady, M. Schaap, B. Spain, H. Kim, C. Taylor, HeartFlow; I. Vignon-Clementel, REO)

In this work, we have proposed a blood-vessel generation procedure for extending known patient vasculature over and within the heart ventricle [8]. It is patient-specific, in the sense that it extend the known, segmented patient vasculature, and it is consistent with physics-based blood vessels characteristics (i.e. derived from CFD) and known vessel physiology. The generated vascular network bridges the gap between the vasculature that can be imaged and assessed via classical means (CT or MRI) and perfusion maps that can be imaged with specific modalities (radiotracer injected scintigraphy or PET). One objective of this work is to eventually propose a forward model for perfusion map generation, that can be used to solved the associated inverse problem of finding the cause of observed perfusion deficits associated with coronary diseases that cannot be imaged directly.

7.19. Curvilinear structure analysis using path operators

Participant: Hugues Talbot (in collaboration with O. Merveille, N. Passat, CRESTIC, and L. Najman, ESIEE Paris)

In this work, we propose mathematical morphology based operators that use paths as families of structuring elements [12]. Structuring elements are like the windows of linear operators, they define the extent of the related operators (convolutions in the linear case, openings and closings in the morphology case). When dealing with thin objects (e.g. fibres, blood vessels, textures, etc), a compact, isotropic window is usually inappropriate because no such window can fit in these objects. This is more critical for morphology, which is concerned with preserving shapes, than with linear operators. Thin windows must therefore be devised, but there are a large number of potentially interesting thin windows at each point in an image. In this article, we leverage the definition of noise-resistant, path operators to define a non-linear notion of vesselness, that can be used for thin object detection, filtering and segmentation in 2D and 3D.

7.20. High throughput automated detection of axial malformations in fish embryo

Participant: Hugues Talbot (in collaboration with D. Genest, M. Léonard, N. De Crozé, L'Oréal, and E. Puybareau, J. Cousty, LIGM)

Fish embryos are used throughout the cosmetics industry to assess the toxicity of the components of their products, as well as more generally in waterways pollution measurements. Indeed pollution is often detectable in trace amounts when they hinder, stop or cause malformations during fish embryo development. In this work, we propose a high-throughput procedure for detecting tail malformation in fish embryo, based on image analysis and machine learning [5]. These malformation are among the most difficult to assess but very common in various degrees of severity. Our procedure provide similar error rate as trained and careful humans operators, as assessed on thousands of images acquired in partneship with L'Oréal. We also show that our procedure is much faster and more consistent than human operators. It is now used in production by our partner.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

PhD Contract with General Electric Healthcare
Project title: Minimally invasive assessement of coronary disease
Duration: 2018-2021
Leader: Hugues Talbot

PhD Contract with General Electric Healthcare
Project title: Optimization methods for breast tomosynthesis
Duration: 2017-2020
Leader: J.-C. Pesquet

PhD Contract with IFP Energies nouvelles
Project title: Graph-based learning from integrated multi-omics and multi-species data
Duration: 2019-2022
Leader: F. Malliaros and J.-C. Pesquet

GPU grant from NVIDIA
NVIDIA's Academic Programs Team is dedicated to empowering and collaborating with professors and researchers at universities worldwide. For a research project on compressing CNNs input, Edouard Oyallon received a TitanXP from NVIDIA.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

Program: ANR PRC

Project acronym: CoMeDIC

Project title: Convergent Metrics for DIscrete Calculus

Duration: 2016-2021

Coordinator: J.-O. Lachaud (Univ. Rhones Alpes Savoie Mont-Blanc), Local: H. Talbot

Program: ANR PRCE

Project acronym: R-Vessel-X

Project title: Extraction et interprétation robustes des réseaux vasculaires dans les images biomédicales hépatiques

Duration: 2018-2022

Coordinator: A. Vacavant (Univ. Clermont Auvergne), local: H. Talbot

Program: ANR JCJC

Project acronym: LearnCost

Project title: Learning Model Constraints for Structured Prediction

Duration: 2014-2018

Coordinator: M. Blaschko

Program: ANR JCJC

Project acronym: MajIC

Project title: Majorization-Minimization Algorithms for Image Computing

Duration: 2017-2021

Coordinator: E. Chouzenoux

Program: ANR JCJC

Project acronym: AVENUE

Project title: A Visual memory network for scene understanding

Duration: 2018-2022

Coordinator: Dr. Karteek Alahari (Inria Grenoble - Rhône-Alpes). Local: F. Malliaros.

9.1.2. Others

Program: CNRS MASTODONS
Projet acronym: TABASCO
Project title: Traitement du bruit non Gaussien en spectroscopie
Duration: 2016-2018
Coordinator: E. Chouzenoux

Program: CNRS-CEFIPRA
Project acronym: NextGenBP
Project title: Looking Beyond Backpropagation in Deep Learning
Duration : 2017-2019
Coordinator: E. Chouzenoux

Program: CNRS MI
Projet acronym: SUPREMA
Project title: Super-résolution en microscopie biphotonique
Duration: 2018
Coordinator: E. Chouzenoux

Program: PHC - Campus France
Projet acronym: POLONIUM
Project title: When Poisson and Gauss meet in imaging
Duration: 2018-2020
Coordinator: E. Chouzenoux

9.2. European Initiatives

9.2.1. H2020 Projects

Program: H2020 ITN Marie Sklodowska-Curie

Project acronym: SUNDIAL

Project title: SUrvey Network for Deep Imaging Analysis and Learning

Duration: 2017-2021

Coordinator: R. Peletier (Univ. Groningen, NL), local: H. Talbot

9.3. International Initiatives

9.3.1. Informal International Partners

Sup'Com Tunis - Pr. Amel Benazza-Benhayia. Collaboration Topic: Multispectral imaging.

University of Patras, Greece - Dr. V. Megalooikonomou. Collaboration Topic: Biosignal analysis.

University of Pennsylvania - Prof. Aristeidis Sotiras. Collaboration Topic: Higher Order Graphs in biomedical image analysis.

University of Montréal, MILA - Dr. Eugene Belilovsky, Pr. Simon Lacoste-Julien. Collaboration Topic : Deep learning, scattering transform.

Berkeley University - Dr. Michael Eickenberg and Dr. Damien Scieur. Collaboration Topic : Deep learning.

KU Leuven - Pr. Matthew Blashcko. Collaboration Topic : Scattering transform.

University of Amsterdam - Dr. Jörn Jacobsen. Collaboration Topic : Deep learning.

Aristotle University of Thessaloniki, Greece - Prof. Apostolos N. Papadopoulos. Collaboration Topic: Graph mining and learning.

Indraprastha Information Institute Technology, Delhi, India - Dr. Angshul Majumdar. Collaboration Topic: Dictionary learning.

Universidad Técnica Federico Santa María, Valparaíso, Chile - Dr. Luis M. Briceño-Arias. Collaboration Topic: Stochastic optimization.

North Carolina State University - Prof. Patrick Louis Combettes. Collaboration Topic: Stochastic optimization.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

Dr. Luis M. Briceño-Arias, Universidad Técnica Federico Santa María, Valparaíso, Chile, 1 Jun. - 1 Jul. 2018, 21 Nov. - 21 Dec. 2018

Jyoti Maggu (PhD student), IIIT New Delhi, India, 05 Mar.-28 May 2018

Vanika Singhal (PhD student), IIIT New Delhi, India, 05 Mar.-20 Apr. 2018

Georgios Panagopoulos (PhD student), Ecole Polytechnique, 15 Jun. - 31 Jul. 2018

9.4.2. Visits to International Teams

9.4.2.1. Research Stays Abroad

M.C. Corbineau, Department of Physics, Informatics and Mathematics, Universita degli studi di Modena e Reggio Emilia, Modena, Italy, 20 Sep. - 20 Oct. 2018.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. Member of the Organizing Committees

R. Güler organized the Workshop “COCO + Mapillary Joint Recognition Challenge” at ECCV 2018, September 2018, Munich, Germany.

R. Güler organized the Workshop “Posetrack Challenge Articulated People Tracking in the Wild”, at ECCV 2018, September 2018, Munich, Germany.

F. Malliaros was a member of the organizing committee of the 3$^{\text{rd}}$ International Workshop on Learning Representations for Big Networks (BigNet), The Web Conference (WWW), Lyon, France, 2018.

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

H. Talbot chaired the 2nd Workshop on Reproducible Research in Pattern Recognition organized at ICPR 2018, Beijing, China.

10.1.2.2. Member of the Conference Program Committees

H. Talbot. Member of the board and program committee for the International Symposium in Mathematical Morphology, to take place in Saarbrücken, Germany, May 2019.

E. Chouzenoux. Member of the technical committees “Signal Processing Theory and Methods” of the IEEE Signal Processing Society and “SAT Signal and Data Analytics for Machine Learning” of EURASIP.

F. Malliaros. Member of the program committee at: AAAI Conference on Artificial Intelligence (AAAI), The Web Conference (WWW), Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL), Conference on Neural Information Processing Systems (NeurIPS), International Conference on Complex Networks and Their Applications (Complex Networks)

10.1.2.3. Reviewer

The members of the team reviewed numerous papers for several international conferences, such as for the annual conferences on Computer Vision and Pattern Recognition (CVPR), Medical Image Computing and Computer Assisted Intervention (MICCAI), Neural Information Processing Systems (NIPS) and International Conference on Learning Representations (ICLR), IEEE International Conference and Acoustics Speech and Signal Processing (ICASSP), IEEE International Conference on Image Processing (ICIP), IEEE Statistical Signal Processing workshop (SSP), European Signal Processing Conference (EUSIPCO), Joint Urban Remote Sensing Event (JURSE).

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

Hugues Talbot: Senior Area Editor for IEEE Signal Processing Letters

Maria Vakalopoulou: Lead Guest Editor for the special issue on Remote Sensing, in Computer Vision and Image Understanding (CVIU)

Emilie Chouzenoux: Associate Editor of the IEEE Transactions on Signal Processing

Jean-Christophe Pesquet: Associate Editor of the SIAM Journal on Imaging Sciences

10.1.3.2. Reviewer - Reviewing Activities

H. Talbot: IEEE Pattern Analysis and Machine Intelligence, IEEE Transaction on Image Processing, Computer Vision and Image Understanding, J. on Mathematical Imaging and Vision, Signal Processing Letters, Transactions on Signal Processing.

M. Vakalopoulou: International Journal of Computer Assisted Radiology and Surgery (IJCARS), IEEE Trans. on Geoscience and Remote Sensing (TGRS), Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS), ISPRS Journal of Photogrammetry and Remote Sensing, Computer Methods and Programs in Biomedicine (CMPB), Pattern Recognition Letters, Computer Vision and Image Understanding (CVIU), SPIE Optical Engineering (OE), Remote Sensing MDPI.

E. Chouzenoux: IEEE Trans. on Image Processing, IEEE Trans. Signal Processing, SIAM Journal on Imaging Science, Journal of Optimization Theory and Applications, Journal of Global Optimization.

J.-C. Pesquet: IEEE Trans. on Signal Processing, IEEE Trans. on Image Processing, IEEE Trans. on Information Theory (IEEE-TI), Signal Processing, SIAM Journal on Optimization, SIAM Journal on Imaging Sciences, Journal of Mathematical Imaging and Vision, Journal of Optimization Theory and Applications, Mathematical Programming.

F. Malliaros: Data Mining and Knowledge Discovery (DAMI), ACM Transactions on Knowledge Discovery from Data (TKDD).

M. C. Corbineau: Signal Processing: Image Communication (SPIC), Computer Vision and Image Understanding (CVIU).

M. Papadomanolaki: Computer Vision and Image Understanding (CVIU), Journal of Basic and Applied Research International, Journal of Geography, Environment and Earth Science International.

M. Sahasrabudhe: Computer Vision and Image Understanding (CVIU).

D. Khuê Lê-Huu: IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), International Journal of Computer Vision (IJCV).

10.1.4. Invited Talks

J.-C. Pesquet:
SAMSI workshop on Operator Splitting Methods in Data Analysis, Raleigh, March 2018,
New York University, March 2018,
Entropy workshop, Barcelona, May 2018,
SIAM Conference on Imaging Science, Bologna, Italy, June 2018,
Aalto University, Helsinki, October 2018,
Polish Science Academy, Warsaw, December 2018.

H. Talbot:
Heartflow, Redwood City, California, April 2018,
Imperial College, London, October 2018,
Institut Henri Poincaré, Paris, December 2018.

E. Chouzenoux:
CRIStAL, Lille, February 2018,
Ann Arbor University, USA, April 2018,
International Conference Inverse Problems: Modeling and Simulation (IPMS 2018), Malta, 21st-25th May 2018,
Heriot Watt University, Edimburg, UK, June 2018,
SIAM Conference on Imaging Science, Bologna, Italy, June 2018.

F. Malliaros:
The Web Conference, Lyon, France, April 2018,
Paris Descartes University, Paris, France, May 2018,
ACM International Conference on Information and Knowledge Management (CIKM), Turin, Italy, October 2018.

E. Oyallon:
SONY CSL, Paris, January 2018,
DeepMind CSML Seminar Series, January 2018,
Imaging in Paris Seminar, IHP, January 2018,
Criteo, Paris, May 2018,
SequeL, Lille, May 2018,
GE Healthcare, Bures-sur-Yvette, June 2018,
NAVER LABS, Grenoble, June 2018,
GT DeepNet, LRI/TAO, Gif-sur-Yvette, December 2018,
GREYC, Caen, December 2018.

M. Vakalapolou:
ONERA, March 2018,
Telecom ParisTech - LTCI, May 2018,
CNRS Workshop on Artificial Intelligence and its Applications, CentraleSupélec, July 2018.

A. Celikkanat:
Paris-Saclay Junior Conference on Data Science and Engineering (JDSE2018) (Poster), Orsay, France, 13-14th September 2018.

M. C. Corbineau:
SIAM Conference on Imaging Science, Bologna, Italy, June 2018.

R. A. Guler:
Inria Paris, March 2018,
University of Amsterdam, April 2018
Max Planck Institute for Intelligent Systems, Tubingen, July 2018.

10.1.5. Leadership within the Scientific Community

J.-C. Pesquet is senior member of the Institut Universitaire de France and a Fellow of IEEE.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master : Corbineau, Marie-Caroline and Pesquet, Jean-Christophe. Advanced course on Optimization, 33h, M1, CentraleSupélec, FR

Master : Chouzenoux, Emilie and Pesquet, Jean-Christophe and Corbineau, Marie-Caroline. Foundations of Distributed and Large Scale Computing, 26h, M.Sc. in Data Sciences and Business Analytics, CentraleSupélec, MVA ENS Cachan, Master Optimization Paris Sud and ESSEC Business School, FR

Master: Pesquet, Jean-Christophe. Introduction to Optimization, 6h, MVA ENS Cachan, FR

Master: Malliaros, Fragkiskos. Machine Learning, 27h, M.Sc. in Data Sciences and Business Analytics, CentraleSupélec and ESSEC Business School, FR

Master: Malliaros, Fragkiskos. Network Science Analytics, 27h, M.Sc. in Data Sciences and Business Analytics, CentraleSupélec and ESSEC Business School and M.Sc. in Artificial Intelligence, CentraleSupélec, FR

Master: Malliaros, Fragkiskos. Introduction to Machine Learning, 33h, 2^{nd} year course at CentraleSupélec, FR

Master: Malliaros, Fragkiskos. Mathematical Modeling of Propagation Phenomena – Propagation on Graphs, 15h, 1^{st} year course at CentraleSupélec, FR

Master: Sahasrabudhe, Mihir: Programming and Languages, 24h, M.Sc. in Data Science and Business Analytics, CentraleSupélec and ESSEC Business School, FR

Master: Sahasrabudhe, Mihir and Battistella, Enzo: Introduction to Machine Learning, 28h, 2nd year CentraleSupélec, FR

Master: Chouzenoux, Emilie. Advanced Machine Learning, 24h, 3rd year CentraleSupélec, FR

Master: Papadomanolaki, Maria. Deep Learning, 3rd year CentraleSupélec and MVA ENS Cachan, FR

Master: Marmin, Arthur, Cours de soutien de mathématiques, 28h, 2nd year CentraleSupélec, FR

Master: Celikkanat, Abdulkadir. Network Science Analytics, 20h, M.Sc. in Data Sciences and Business Analytics, ESSEC and 3rd year CentraleSupélec, FR

Master: Celikkanat, Abdulkadir. Mathematical Modeling of Propagation Phenomena - Propagation on Graphs, 12h, 3rd year CentraleSupélec, FR

Master: Estienne, Théo. Deep Learning, 7h30, 3rd year CentraleSupélec, FR

Master: Papadomanolaki, Maria. Deep Learning, 7h30, 3rd year CentraleSupélec, FR

Master: Oyallon, Edouard. Deep Learning, 12h, 3rd year CentraleSupélec, FR

Master: Oyallon, Edouard. Deep Learning, 1h, MVA ENS Cachan, FR

Master: Vakalopoulou, Maria. Deep Learning, 12h, 3rd year CentraleSupélec, FR

Master: Vakalopoulou, Maria. Introduction to Machine Learning, 12h, 2nd year CentraleSupélec, FR

Master: Talbot, Hugues. Discrete Optimisation, 2nd year course, CentraleSupelec, 30h, FR

Master: Talbot, Hugues. Big Data, Techniques and Platforms, M.Sc in Data Science and Business Analytics, CentraleSupelec and ESSEC Business School, 30h, FR

10.2.2. Lecturing activities

Chouzenoux, Emilie and Talbot, Hugues. Enjeux et technologies de l'intelligence artificielle. CentraleSupélec Exed, 2 days ($\times 2$), FR.

Oyallon, Edouard. Introduction to Deep Learning. Ateliers Statistiques de la Société Française de Statistique, 2 days, IHP, Paris, FR.

Oyallon, Edouard. Advanced Deep Learning. Mathematical Coffees, 3 hours, Huawei, Paris, FR.

Chouzenoux, Emilie. Parallel Stochastic Computing. Mathematical Coffees, 3 hours, Huawei, Paris, FR.

Vakalopoulou, Maria. Introduction to Computer Vision, Applications to Medical Images. Summer School on Artificial Intelligence, 5 hours, CentraleSupélec, FR.

Pesquet, Jean-Christophe. Proximal Splitting Methods in Image Processing, Lake Como School of Advanced Studies on Computational Methods for Inverse Problems in Imaging, 12 hours, May 2018, Italy.

10.2.3. Supervision

PhD (defended) : Siddhartha Chandra, Deep Structured Prediction for Dense Labeling Tasks in Computer Vision, 2014-2018, supervised by Iasonas Kokkinos. Defended on 11st May 2018.

PhD (defended) : Stefan Kinauer, Représentations à base de parties pour la vision 3D de haut niveau, 2014-2018, supervised by Iosinas Kokkinos. Defended on 31st August 2018.

PhD (defended) : Afef Cherni, Méthodes modernes d'analyse de données en biophysique analytique, 2015-2018, supervised by Emilie Chouzenoux and Marc-André Delsuc (IGBMC, Strasbourg). Defended on 20th September 2018.

PhD (defended): Clara Jaquet. Vers la simulation de perfusion du myocarde à partir d'image tomographique scanner. 2014-2018. Supervised by Hugues Talbot and Laurent Najman (ESIEE). Defended on 13 December 2018.

PhD in progress : Mihir Sahasrabudhe, Understanding Correlations in High-Dimensional Spaces and their Applications in Medical Imaging and Computer Vision, 2015-2019, supervised by Nikos Paragios

PhD in progress : Marie-Caroline Corbineau, Fast online optimization algorithms for machine learning and medical imaging, 2016-2019, supervised by Emilie Chouzenoux and J.-C. Pesquet

PhD in progress : Loubna El Gueddari, Parallel proximal algorithms for compressed sensing MRI reconstruction - Applications to ultra-high magnetic field imaging, 2016-2019, supervised by J.-C. Pesquet and Ph. Ciuciu (Inria PARIETAL)

PhD in progress: Diane Genest. Imagerie du Modele alevin de poisson / Application a la toxicologie du developpement. 2016-2019. Supervised by Hugues Talbot and Jean Cousty (ESIEE).

PhD in progress: Daniel Antunes: Contraintes géométriques et approches variationnelles pour l'analyse d'image. 2016-2019. Supervised by Hugues Talbot and Jacques-Olivier Lachaud (U. Savoie-Mont Blanc)

PhD in progress: Guillaume Chassagnon, Development of new quantitative imaging biomarkers for obstructive and interstitial lung diseases, 2016-2019, supervised by N. Paragios

PhD in progress: Maïssa Sghaier, clinical Task-Based Reconstruction in tomosynthesis, 2017-2020, supervised by E. Chouzenoux, J.-C. Pesquet and G. Palma (GE Healthcare)

PhD in progress: Arthur Marmin, Rational models optimized exactly for chemical processes improvement, 2017-2020, supervised by Marc Castella (Telecom Paristech) and J.-C. Pesquet

PhD in progress: Roger Sun, Deep learning and computer vision approaches on medical imaging and genomic data to improve the prediction of anticancer therapies' efficacy, 2017-2020, supervised by N. Paragios

PhD in progress: Marie-Charlotte Poilpre: Méthode de comparaison faciale morphologique, adaptée aux expertise judiciaires, basée sur la modélisation 3D. 2017-2020. Supervised by Hugues Talbot and Vincent Nozick (U. Paris-Est)

PhD in progress: Théo Estienne, Improving anticancer therapies efficacy through Machine Learning on Medical Imaging & Genomic Data, 2017-2020, supervised by N. Paragios

PhD in progress: Abdulkadir Celikkanat, Representation learning methods on graphs, 2017-2020, supervised by N. Paragios (TheraPanacea, Paris) and F. Malliaros

PhD in progress: Maria Papadomanolaki, Change Detection from Multitemporal High Resolution Data with Deep Learning, 2017-2021, supervised by M. Vakalopoulou and K. Karantzalos (National Technical University of Athens).

PhD in progress: Enzo Battistella, Development of novel imaging approaches for tumour phenotype assessment by noninvasive imaging 2017-2020, supervised by M. Vakalopoulou and N. Paragios.

PhD in progress: Thank Xuan Nguyen. Détection et étude morphologique des sources extra-galactiques par analyse variationnelle. 2018-2021. Supervised by Hugues Talbot and Laurent Najman (ESIEE)

PhD in progress: Marvin Lerousseau. Apprentissage statistique en imagerie médicale et en génomique pour prédire l'efficacité des thérapies anti-tumorales. 2018-2021. Supervised by Nikos Paragios (Therapanacea), Eric Deutch (IGR) and Hugues Talbot.

PhD in progress: Yunshi Huang, Majorization-Minimization approaches for large scale problems in image processing, 2018-2021, supervised by E. Chouzenoux.

PhD in progress: Georgios Panagopoulos, Influence maximization in social networks, 2018-2021, supervised by F. Malliaros and M. Vazirgiannis (Ècole Polytechnique).

10.2.4. Juries

The faculty members of the team participated to numerous PhD Thesis Committees, HDR Committees and served as Grant Reviewers.
Emilie Chouzenoux, Marie-Caroline Corbineau, Fragkiskos Malliaros, Edouard Oyallon, Jean-Christophe Pesquet, Mihir Sahasrabudhe and Hugues Talbot were part of the jury committee for several end-of-course internship presentations in CentraleSupélec.

10.3. Popularization

10.3.1. Interventions

Enzo Batistella, Marie-Caroline Corbineau, Théo Estienne and Maria Vakalopoulou have participated at the 'Demi-journée Recherche' organised by CentraleSupélec for the first year students of the school.

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] S. CHANDRA.*Efficient Deep Structured Prediction for Dense Labeling Tasks in Computer Vision*, Université Paris-Saclay, May 2018, https://tel.archives-ouvertes.fr/tel-01812763

[2] S. KINAUER.*Part-Based Representations for High-Level 3D Vision*, Université Paris-Saclay, August 2018, https://tel.archives-ouvertes.fr/tel-01885958

Articles in International Peer-Reviewed Journal

[3] F. ABBOUD, E. CHOUZENOUX, J.-C. PESQUET, J.-H. CHENOT, L. LABORELLI.*An Alternating Proximal Approach for Blind Video Deconvolution*, in "Signal Processing: Image Communication", 2018, vol. 70, p. 21-36, https://hal.archives-ouvertes.fr/hal-01668437

[4] J. CASAS-ROMA, J. SALAS, F. MALLIAROS, M. VAZIRGIANNIS.*k-Degree anonymity on directed networks*, in "Knowledge and Information Systems (KAIS)", September 2018 [*DOI :* 10.1007/s10115-018-1251-5], https://hal-centralesupelec.archives-ouvertes.fr/hal-01950285

[5] D. GENEST, É. PUYBAREAU, M. LÉONARD, N. DE CROZÉ, J. COUSTY, H. TALBOT.*High throughput automated detection of axial malformations in Medaka fish embryo*, in "Computers in Biology and Medicine", 2019, https://hal.archives-ouvertes.fr/hal-01971148

[6] M. E. GHECHE, G. CHIERCHIA, J.-C. PESQUET.*Proximity Operators of Discrete Information Divergences*, in "IEEE Transactions on Information Theory", February 2018, vol. 64, n° 2, p. 1092-1104, https://hal.archives-ouvertes.fr/hal-01672646

[7] E. GROSSIORD, B. NAEGEL, H. TALBOT, L. NAJMAN, N. PASSAT.*Shape-based analysis on component-graphs for multivalued image processing*, in "Mathematical Morphology - Theory and Applications", 2019, https://hal.univ-reims.fr/hal-01695384

[8] C. JAQUET, L. NAJMAN, H. TALBOT, L. GRADY, M. SCHAAP, B. SPAIN, H. J. KIM, I. VIGNON-CLEMENTEL, C. A. TAYLOR.*Generation of patient-specific cardiac vascular networks: a hybrid image-based and synthetic geometric model*, in "IEEE Transactions on Biomedical Engineering", 2018 [*DOI :* 10.1109/TBME.2018.2865667], https://hal.archives-ouvertes.fr/hal-01869264

[9] C. KARAKIZI, K. KARANTZALOS, M. VAKALOPOULOU, G. ANTONIOU.*Detailed Land Cover Mapping from Multitemporal Landsat-8 Data of Different Cloud Cover*, in "Remote Sensing", August 2018, https://hal.inria.fr/hal-01959065

[10] P. KONDAXAKIS, K. GULZAR, S. KINAUER, I. KOKKINOS, V. KYRKI.*Robot–Robot Gesturing for Anchoring Representations*, in "IEEE Transactions on Robotics", October 2018, p. 1-15, https://hal.archives-ouvertes.fr/hal-01961433

[11] Y. MARNISSI, E. CHOUZENOUX, A. BENAZZA-BENYAHIA, J.-C. PESQUET.*An Auxiliary Variable Method for Markov Chain Monte Carlo Algorithms in High Dimension*, in "Entropy", 2018, vol. 20, n° 2 [*DOI :* 10.3390/E20020110], https://hal.archives-ouvertes.fr/hal-01797093

[12] O. MERVEILLE, H. TALBOT, L. NAJMAN, N. PASSAT.*Curvilinear structure analysis by ranking the orientation responses of path operators*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", 2018, vol. 40, n° 2, p. 304-317 [*DOI :* 10.1109/TPAMI.2017.2672972], https://hal.archives-ouvertes.fr/hal-01262728

[13] E. OYALLON, S. ZAGORUYKO, G. HUANG, N. KOMODAKIS, S. LACOSTE-JULIEN, M. BLASCHKO, E. BELILOVSKY.*Scattering Networks for Hybrid Representation Learning*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", September 2018, 11, https://arxiv.org/abs/1809.06367 [*DOI :* 10.1109/TPAMI.2018.2855738], https://hal.inria.fr/hal-01837587

[14] A. PIRAYRE, C. COUPRIE, L. DUVAL, J.-C. PESQUET.*BRANE Clust: Cluster-Assisted Gene Regulatory Network Inference Refinement*, in "IEEE/ACM Transactions on Computational Biology and Bioinformatics", June 2018, vol. 15, n° 3, p. 850-860, https://hal-ifp.archives-ouvertes.fr/hal-01330638

[15] M.-E. Rossi, B. Shi, N. Tziortziotis, F. Malliaros, C. Giatsidis, M. Vazirgiannis.*MATI: An efficient algorithm for influence maximization in social networks*, in "PLoS ONE", October 2018, https://hal-centralesupelec.archives-ouvertes.fr/hal-01950273

[16] P. Savadjiev, J. Chong, A. Dohan, M. Vakalopoulou, C. Reinhold, N. Paragios, B. Gallix.*Demystification of AI-driven medical image interpretation: past, present and future*, in "European Radiology", August 2018, https://hal.inria.fr/hal-01958231

[17] R. Sun, E. J. Limkin, M. Vakalopoulou, L. Dercle, S. Champiat, S. R. Han, L. Verlingue, D. Brandao, A. Lancia, S. Ammari, A. Hollebecque, J.-Y. Scoazec, A. Marabelle, C. Massard, J.-C. Soria, C. Robert, N. Paragios, É. Deutsch, C. Ferté.*A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study*, in "The Lancet Oncology", September 2018, https://hal.inria.fr/hal-01958243

Invited Conferences

[18] P. L. Combettes, J.-C. Pesquet.*Linear convergence of stochastic block-coordinate fixed point algorithms*, in "European Signal and Image Processing Conference (EUSIPCO 2018)", Rome, Italy, September 2018, https://hal.archives-ouvertes.fr/hal-01964580

[19] M.-C. Corbineau, E. Chouzenoux, J.-C. Pesquet.*Geometry-Texture Decomposition/Reconstruction Using a Proximal Interior Point Algorithm*, in "10th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2018)", Sheffield, United Kingdom, Proceedings of the 10th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2018), July 2018, https://hal.archives-ouvertes.fr/hal-01863408

[20] J. Maggu, A. Majumdar, E. Chouzenoux.*Transformed Locally Linear Manifold Clustering*, in "26th European Signal Processing Conference", Rome, Italy, Proceedings of the 26th European Signal Processing Conference (EUSIPCO 2018), September 2018, https://hal.archives-ouvertes.fr/hal-01862192

International Conferences with Proceedings

[21] F. Abboud, E. Chouzenoux, J.-C. Pesquet, H. Talbot.*A Multicore Convex Optimization Algorithm with Applications to Video Restoration*, in "IEEE International Conference on Image Processing", Athens, Greece, Proceedings of the IEEE International Conference on Image Processing (ICIP 2018), October 2018, https://hal.archives-ouvertes.fr/hal-01862210

[22] A. Benfenati, E. Chouzenoux, J.-C. Pesquet.*A Nonconvex Variational Approach for Robust Graphical Lasso.*, in "IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2018)", Calgary, Canada, Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2018), April 2018, https://hal.archives-ouvertes.fr/hal-01803424

[23] S. Chandra, M. Vakalopoulou, L. Fidon, E. Battistella, T. Estienne, R. Sun, C. Robert, É. Deutsch, N. Paragios.*Context Aware 3D CNNs for Brain Tumor Segmentation*, in "MICCAI Brainlesion Workshop", Granada, Spain, September 2018, https://hal.inria.fr/hal-01959610

[24] A. Cherni, E. Chouzenoux, D. Marc-André.*Fast Dictionnary-Based Approach for Mass Spectrometry Data Analysis*, in "IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2018)", Calgary, Canada, Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2018), April 2018, https://hal.archives-ouvertes.fr/hal-01803419

[25] S. CHRISTODOULIDIS, M. SAHASRABUDHE, M. VAKALOPOULOU, G. CHASSAGNON, M.-P. REVEL, S. MOUGIAKAKOU, N. PARAGIOS.*Linear and Deformable Image Registration with 3D Convolutional Neural Networks*, in "Reconstruction and Analysis of Moving Body Organs, 21th International Conference on Medical Image Computing and Computer Assisted Intervention 2018", Grenada, Spain, September 2018, https://hal.archives-ouvertes.fr/hal-01935607

[26] M.-C. CORBINEAU, E. CHOUZENOUX, J.-C. PESQUET.*PIPA: A New Proximal Interior Point Algorithm for Large Scale Convex Optimization*, in "IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2018)", Calgary, Canada, Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2018), April 2018, https://hal.archives-ouvertes.fr/hal-01803422

[27] R. A. GÜLER, N. NEVEROVA, I. KOKKINOS.*DensePose: Dense Human Pose Estimation In The Wild*, in "Conference on Computer Vision and Pattern Recognition (CVPR) 2018", Salt Lake City, United States, June 2018, https://hal.archives-ouvertes.fr/hal-01951864

[28] J.-H. JACOBSEN, A. SMEULDERS, E. OYALLON.*i-RevNet: Deep Invertible Networks*, in "ICLR 2018 - International Conference on Learning Representations", Vancouver, Canada, April 2018, https://arxiv.org/abs/1802.07088 , https://hal.archives-ouvertes.fr/hal-01712808

[29] D. K. LÊ-HUU, N. PARAGIOS.*Continuous Relaxation of MAP Inference: A Nonconvex Perspective*, in "CVPR 2018 - IEEE Conference on Computer Vision and Pattern Recognition", Salt Lake City, United States, June 2018, p. 1-19, https://hal.inria.fr/hal-01716514

[30] J. MAGGU, E. CHOUZENOUX, G. CHIERCHIA, A. MAJUMDAR.*Convolutional Transform learning*, in "25th International Conference on Neural Information Processing (ICONIP 2018)", Siem Reap, Cambodia, Lecture Notes in Computer Science, December 2018, vol. 11303, p. 162-174, https://hal.archives-ouvertes.fr/hal-01862201

[31] A. MARMIN, M. CASTELLA, J.-C. PESQUET, L. DUVAL.*Signal Reconstruction from Sub-sampled and Nonlinearly Distorted Observations*, in "26th European Signal Processing Conference (EUSIPCO 2018)", Roma, Italy, September 2018, https://hal.archives-ouvertes.fr/hal-01957568

[32] N. NEVEROVA, R. A. GÜLER, I. KOKKINOS.*Dense Pose Transfer*, in "European Conference on Computer Vision (ECCV) 2018", Munich, Germany, September 2018, https://hal.archives-ouvertes.fr/hal-01951850

[33] D. NGUYEN, F. MALLIAROS.*BiasedWalk: Biased Sampling for Representation Learning on Graphs*, in "International Workshop on High Performance Big Graph Data Management, Analysis, and Mining (BigGraphs)", Seattle, WA, United States, December 2018, https://hal-centralesupelec.archives-ouvertes.fr/hal-01958902

[34] E. OYALLON, E. BELILOVSKY, S. ZAGORUYKO, M. VALKO.*Compressing the Input for CNNs with the First-Order Scattering Transform*, in "European Conference on Computer Vision", Munich, Germany, 2018, https://hal.inria.fr/hal-01850921

[35] G. PANAGOPOULOS, F. MALLIAROS, M. VAZIRGIANNIS.*DiffuGreedy: An Influence Maximization Algorithm based on Diffusion Cascades*, in "The 7th International Conference on Complex Networks and Their Applications (Complex Networks)", Cambridge, United Kingdom, December 2018, https://hal-centralesupelec.archives-ouvertes.fr/hal-01958915

[36] M. PAPADOMANOLAKI, M. VAKALOPOULOU, N. PARAGIOS, K. KARANTZALOS.*Stacked Encoder-Decoders for Accurate Semantic Segmentation of Very High Resolution Satellite Datasets*, in "IGARSS 2018 - 38th annual International Geoscience and Remote Sensing Symposium", Valencia, Spain, July 2018, p. 1-4, https://hal.archives-ouvertes.fr/hal-01870857

[37] B. PASCAL, N. PUSTELNIK, P. ABRY, J.-C. PESQUET.*Block-coordinate proximal algorithms for scale-free texture segmentation*, in "ICASSP 2018 - IEEE International Conference on Acoustics, Speech and Signal Processing", Calgary, Canada, April 2018, p. 1-5, https://hal.inria.fr/hal-01736991

[38] Z. SHU, M. SAHASRABUDHE, R. A. GÜLER, D. SAMARAS, N. PARAGIOS, I. KOKKINOS.*Deforming Autoencoders: Unsupervised Disentangling of Shape and Appearance*, in "The European Conference on Computer Vision", Munich, Germany, September 2018, https://hal.archives-ouvertes.fr/hal-01935596

[39] K. SKIANIS, F. MALLIAROS, M. VAZIRGIANNIS.*Fusing Document, Collection and Label Graph-based Representations with Word Embeddings for Text Classification*, in "NAACL-HLT Workshop on Graph-Based Natural Language Processing (TextGraphs)", New Orleans, Louisiana, United States, June 2018, https://hal-centralesupelec.archives-ouvertes.fr/hal-01848880

[40] T. TSZ-KIT, E. CHOUZENOUX, C. LEFORT, J.-C. PESQUET.*Optimal Multivariate Gaussian Fitting for PSF Modeling in Two-Photon Microscopy*, in "15 th IEEE International Symposium on Biomedical Imaging (ISBI 2018)", Washington DC, United States, 15 th Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI 2018), April 2018 [*DOI :* 10.1109/ISBI.2018.8363621], https://hal.archives-ouvertes.fr/hal-01803413

[41] M. VAKALOPOULOU, G. CHASSAGNON, N. BUS, R. MARINI SILVA, E. I. ZACHARAKI, M.-P. REVEL, N. PARAGIOS.*AtlasNet: Multi-atlas Non-linear Deep Networks for Medical Image Segmentation*, in "International Conference on Medical Image Computing and Computer-Assisted Intervention", Granada, Spain, September 2018, https://hal.inria.fr/hal-01958236

[42] M. VAKALOPOULOU, G. CHASSAGNON, N. PARAGIOS, M.-P. REVEL, E. I. ZACHARAKI.*Deep patch-based priors under a fully convolutional encoder-decoder architecture for interstitial lung disease segmentation*, in "ISBI'18 - IEEE International Symposium on Biomedical Imaging", Washington, D.C., United States, April 2018, https://hal.inria.fr/hal-01721714

[43] M. VAZIRGIANNIS, F. MALLIAROS, G. NIKOLENTZOS.*GraphRep: Boosting Text Mining, NLP and Information Retrieval with Graphs*, in "27th ACM International Conference on Information and Knowledge Management (CIKM)", Torino, Italy, October 2018, p. 2295-2296 [*DOI :* 10.1145/3269206.3274273], https://hal-centralesupelec.archives-ouvertes.fr/hal-01958927

[44] N. ZEGHIDOUR, N. USUNIER, I. KOKKINOS, T. SCHATZ, G. SYNNAEVE, E. DUPOUX.*Learning Filterbanks from Raw Speech for Phoneme Recognition*, in "ICASSP 2018", Calgary, Alberta, Canada, Proceedings of ICASSP 2018, April 2018, https://arxiv.org/abs/1711.01161v2 - Accepted at ICASSP 2018, https://hal.archives-ouvertes.fr/hal-01888737

[45] A. ÇELIKKANAT, F. MALLIAROS.*TNE: A Latent Model for Representation Learning on Networks*, in "32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Relational Representation Learning Workshop", Montréal, Canada, December 2018, https://hal.archives-ouvertes.fr/hal-01957684

Conferences without Proceedings

[46] E. GROSSIORD, H. TALBOT, N. PASSAT, M. MEIGNAN, L. NAJMAN.*Segmentation 3D des lésions du lymphome à partir de descripteurs multimodaux TEP/TDM*, in "Journée thématique du GdR ISIS : "Segmentation d'images biomédicales : quels outils pour l'analyse des données massives, hétérogènes et multimodales ?"", Paris, France, 2018, https://hal.univ-reims.fr/hal-01745773

[47] D. SCIEUR, E. OYALLON, A. D'ASPREMONT, F. BACH.*Nonlinear Acceleration of CNNs*, in "ICLR Workshop track", Vancouver, Canada, April 2018, https://hal.archives-ouvertes.fr/hal-01805251

Research Reports

[48] A. BENFENATI, E. CHOUZENOUX, L. DUVAL, J.-C. PESQUET, A. PIRAYRE.*A review on graph optimization and algorithmic frameworks*, LIGM - Laboratoire d'Informatique Gaspard-Monge, October 2018, https://hal.archives-ouvertes.fr/hal-01901499

[49] C. BERTOCCHI, E. CHOUZENOUX, M.-C. CORBINEAU, J.-C. PESQUET, M. PRATO.*Deep Unfolding of a Proximal Interior Point Method for Image Restoration*, CVN, CentraleSupélec, Université Paris-Saclay, Gif-Sur-Yvette, France ; Universita di Modena e Reggio Emilia, Modena, Italy ; Université Paris-Est Marne la Vallée, LIGM, UMR CNRS 8049, Champs-sur-Marne, France, December 2018, https://arxiv.org/abs/1812.04276 , https://hal.archives-ouvertes.fr/hal-01943475

Other Publications

[50] F. ABBOUD, E. CHOUZENOUX, J.-C. PESQUET, H. TALBOT.*Distributed Algorithms for Proximity Operator Computation with Applications to Video Processing*, December 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01942710

[51] Ö. D. AKYILDIZ, E. CHOUZENOUX, V. ELVIRA, J. MÍGUEZ.*A probabilistic incremental proximal gradient method*, December 2018, https://arxiv.org/abs/1812.01655 - 5 pages, https://hal.archives-ouvertes.fr/hal-01946642

[52] S. BAKAS, M. REYES, E. BATTISTELLA, S. CHANDRA, T. ESTIENNE, L. FIDON, M. VAKALOPOULOU, R. SUN, É. DEUTSCH, C. ROBERT, N. PARAGIOS, B. MENZE.*Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01959023

[53] M. CASTELLA, J.-C. PESQUET, A. MARMIN.*Rational Optimization for Nonlinear Reconstruction with Approximate $\ell 0$ Penalization*, August 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01852289

[54] P. L. COMBETTES, J.-C. PESQUET.*Stochastic Quasi-Fejér Block-Coordinate Fixed Point Iterations With Random Sweeping II: Mean-Square and Linear Convergence*, December 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01964582

[55] A. KANJ, H. TALBOT, J.-C. PESQUET, R. R. LUPARELLO.*A Variational Method for Flicker Removal in High Speed Video Sequences*, March 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01729639

[56] Y. MARNISSI, E. CHOUZENOUX, A. BENAZZA-BENYAHIA, J.-C. PESQUET.*Majorize-Minimize Adapted Metropolis-Hastings Algorithm*, October 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01909153

[57] O. MERVEILLE, B. NAEGEL, H. TALBOT, N. PASSAT.*nD variational restoration of curvilinear structures with directional regularization*, July 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01832636

[58] A. J.-P. TIXIER, M.-E. ROSSI, F. MALLIAROS, J. READ, M. VAZIRGIANNIS.*Perturb and Combine to Identify Influential Spreaders in Real-World Networks*, December 2018, working paper or preprint, https://hal-centralesupelec.archives-ouvertes.fr/hal-01958973

Project-Team GAMMA3

Automatic mesh generation and advanced methods

IN PARTNERSHIP WITH:
Université de Technologie de Troyes

RESEARCH CENTER
Saclay - Île-de-France

THEME
Numerical schemes and simulations

Table of contents

Project-Team GAMMA3

Creation of the Project-Team: 2010 January 01

Keywords:

Computer Science and Digital Science:

A2.5. - Software engineering
A5.2. - Data visualization
A5.5.1. - Geometrical modeling
A6.1. - Methods in mathematical modeling
A6.2. - Scientific computing, Numerical Analysis & Optimization
A7.1. - Algorithms
A8.3. - Geometry, Topology

Other Research Topics and Application Domains:

B5.2.3. - Aviation
B5.2.4. - Aerospace

1. Team, Visitors, External Collaborators

Research Scientists

Paul-Louis George [Team leader, Inria, Senior Researcher]
Frederic Alauzet [Inria, Senior Researcher, HDR]
Patrick Laug [Inria, Senior Researcher, HDR]
Adrien Loseille [Inria, Researcher]
David Marcum [Inria, International Chair, Advanced Research Position]

Faculty Member

Houman Borouchaki [Univ de technologie de Troyes, Professor]

External Collaborators

Julia Camargo [Visiting PhD Student, Stanford University, from Oct 2018 until December 2018]
Eléonore Gauci [Univ Pierre et Marie Curie]
David Marcum [Professor, Mississippi State University]
Loic Marechal [Research Engineer, Distene]

Technical Staff

Victorien Menier [Inria, from Oct 2018, granted by SAFRAN SA]

PhD Students

Bastien Andrieu [ONERA]
Rémi Feuillet [École Nationale Supérieure de Techniques Avancées]
Loïc Frazza [Ecole polytechnique]
Lucille Marie Tenkes [Inria, from Oct 2018]

Post-Doctoral Fellow

Julien Vanharen [Inria]

Visiting Scientist

Nicolas Ringue [MacGill Univiersity, until Jul 2018]

Administrative Assistants

Jessica Gameiro [Inria, until Apr 2018]
Maria Agustina Ronco [Inria, from May 2018]

2. Overall Objectives

2.1. Introduction

Un domaine important des sciences de l'ingénieur concerne le calcul des solutions d'équations aux dérivées partielles très variées (en mécanique du solide, en mécanique des fluides, en modélisation de problèmes thermiques, ...) par la méthode des éléments ou des volumes finis. Ces méthodes utilisent comme support spatial des calculs un maillage du domaine sur lequel les équations sont formulées. Par suite, les algorithmes (de construction) de maillages occupent un rôle primordial dans toute simulation par ces méthodes d'un problème modélisé en équations aux dérivées partielles. En particulier, la précision, voire la validité, des solutions calculées est liée aux propriétés du maillage utilisé [28].

L'équipe-projet GAMMA3 a été créé en 2010 à la suite du projet GAMMA. L'équipe est bilocalisée avec une partie à l'UTT (Troyes) et l'autre à Rocquencourt puis Saclay. Les thèmes du projet regroupent un ensemble d'activités concernant les points indiqués ci-dessus, en particulier, l'aspect génération automatique de maillages afin de construire les supports utilisés par les méthodes d'éléments ou de volumes finis. Sont également étudiés les aspects de modélisation géométrique, de post-traitement et de visualisation des résultats issus de tels calculs [29].

L'évolution de la demande en termes de génération automatique de maillages nécessite une évolution des méthodes classiques de création de maillages vers des méthodes permettant de construire des maillages contrôlés. Les maillages doivent donc être soit isotropes, le contrôle portant sur des tailles souhaitées, soit anisotropes, le contrôle portant à la fois sur des directions et des tailles selon ces dernières.

Le développement d'algorithmes de maillages gouvernés sert de support naturel à la conception de boucles de maillages adaptatifs qui, via un estimateur d'erreurs *a posteriori*, permettent de contrôler la qualité des solutions. Les estimateurs d'erreurs sont issus d'applications en mécanique des fluides (Inria) et du solide (UTT). Leurs validations reposent sur le développement de solveurs avancés, en particulier, en mécanique des fluides. Ces deux points (estimateurs et solveurs) constituent au moins la moitié de nos recherches.

Ces préoccupations amènent à considérer le problème du maillage des domaines de calculs en eux-mêmes tout comme celui du maillage ou du remaillage des courbes et surfaces, frontières de ces domaines.

La taille, en termes de nombre de nœuds, des maillages nécessaires pour certaines simulations, amène à travailler sur la parallélisation des processus de calculs. Cette problématique conduit également à s'intéresser à l'aspect multi-cœurs au niveau des algorithmes de maillages proprement dits.

Simultanément, le volume des résultats obtenus dans de telles simulations, nécessite d'envisager le post-traitement de ces résultats en parallèle ou par des méthodes appropriées.

Par ailleurs, de nombreux problèmes partent de saisies *scanner* (ou autre système discret) des géométries à traiter et demandent d'en déduire des maillages de surfaces aptes à être, par la suite, traités par les méthodes classiques (de remaillage, d'optimisation, de calculs). Cette question, en général mal traitée, reste pertinente.

Enfin, la maturité de certaines méthodes (victimes de leur succès) conduit les utilisateurs à demander plus et à considérer des problèmes de maillage ou des conditions d'utilisations extrêmes induisant des algorithmes *a priori* inattendus.

Les objectifs du projet GAMMA3 consistent à étudier l'ensemble des points mentionnés ci-dessus afin de rendre automatique le calcul de la solution d'un problème donné avec une précision imposée au départ. Par ailleurs, certaines des techniques utilisées dans les problématiques de maillage sont utilisables dans d'autres disciplines (compression d'images pour ne citer qu'un seul exemple).

3. Highlights of the Year

3.1. Highlights of the Year

3.1.1. Awards

- Adrien Loseille. Deuxième Prix FIEEC de la Recherche Appliquée.

4. New Software and Platforms

4.1. ABL4FLO

Adaptive Boundary Layer 4 FLOw
KEYWORDS: Boundary layers - Hybrid meshes
FUNCTIONAL DESCRIPTION: ABL4FLO is a module used to perform adaptive boundary layer mesh adaptation as required in RANS solutions. It is included in Feflo.a/AMG-Lib software. It is based on a constrained version of the cavity operators in order to generate automatically hybrid elements. If a metric surface is provided, the normal and tangential direction are simultaneously adapted.

- Participant: Adrien Loseille
- Contact: Adrien Loseille
- Publications: Recent Improvements on Cavity-Based Operators for RANS Mesh Adaptation - Unstructured Mesh Generation and Adaptation - Robust Boundary Layer Mesh Generation
- URL: https://pyamg.saclay.inria.fr/

4.2. AMA4FLO

Anisotropic Mesh Adaptation 4 FLOw
KEYWORDS: 3D - Mesh adaptation
FUNCTIONAL DESCRIPTION: AMA4Flo is part of Feflo.a which is a robust anisotropic local remeshing software. It is intended for scientific computing with primary applications in aerodynamics and spatial studies. Surface and volume mesh adaptation are handled in a coupled-way. It also includes : - Boundary layers mesh generation for RANS simulations, - CAD re-projection and discrete surface remeshing, with - Hybrid mesh generation for boundary-layers - High-quality quasi-structured grids for complex geometries and complex corners: multi-normals, normals deactivation, ... - Highly anisotropic mesh adaptation, ratios up to 1 million are handled - Anisotropic/Boundary-layer coupling for supersonic shock/boundary layer interaction

The boundary layer module alone (abl4flo) is registered with the APP under nbr. IDDN. FR.001. 080032. 00.S.P.2012. 000.10000

initially, AMA4FLO was mainly focused on Computational Fluid Dynamics (4 FLOw), but it is now sed in many applications areas: seismic, reservoir engineering, spatial, hydrodynamics, hemodynamics, ...

- Participant: Adrien Loseille
- Contact: Adrien Loseille
- Publications: Very High Order Anisotropic Metric-Based Mesh Adaptation in 3D - Computational and Experimental Assessment of Models for the First AIAA Sonic Boom Prediction Workshop Using Adaptive High Fidelity CFD methods - Unique cavity-based operator and hierarchical domain partitioning for fast parallel generation of anisotropic meshes - Unstructured Mesh Generation and Adaptation - A Decade of Progress on Anisotropic Mesh Adaptation for Computational Fluid Dynamics - Metric-orthogonal Anisotropic Mesh Generation - Sonic Boom Assessment of a Hypersonic Transport Vehicle with Advanced Numerical Methods
- URL: https://pyamg.saclay.inria.fr/

4.3. BL2D

KEYWORDS: Abstraction - Meshing - Isotropic - Anisotropic - Delaunay - Mesher - Mesh

FUNCTIONAL DESCRIPTION: This software package stems from a former one called BL2D-V1. The meshing method is of controled Delaunay type, isotropic or anisotropic. The internal point generation follows a frontal logic, and their connection is realised as in a classical Delaunay approach. Quadrilaterals are obtained by a pairing process. The direct construction of degree 2 element has been made possible via the control of the domain boundary mesh, in order to ensure the desired compatibility. The boundary middle nodes are located according to the curvilinear abscissa. The internal middle nodes are, by default, at the middle of the corresponding edges.

RELEASE FUNCTIONAL DESCRIPTION: Par rapport à la version V1, il offre de nombreuses possibilités nouvelles : méthode frontale, triangles quadratiques courbes, quadrilatères de degré 1ou 2, frontières déformables, allocation dynamique de mémoire, etc

- Participants: Houman Borouchaki and Patrick Laug
- Contact: Patrick Laug
- URL: http://pages.saclay.inria.fr/patrick.laug/logiciels/logiciels.html

4.4. BL2D-ABAQ

KEYWORDS: Anisotropic - Delaunay - Automatic mesher - Meshing - Mesher - Mesh
FUNCTIONAL DESCRIPTION: The meshing method is the same as BL2D in an adaptive process. An a posteriori error estimation of a solution at the nodes of the current mesh results in a size map. A new mesh staisfying these size specifications (made continuous is built, and the solution is interpolated on the new mesh.

- Participants: Abel Cherouat, Houman Borouchaki and Patrick Laug
- Contact: Patrick Laug
- URL: http://pages.saclay.inria.fr/patrick.laug/logiciels/logiciels.html

4.5. BLGEOL

KEYWORDS: Automatic mesher - Geologic structure - Meshing - Mesher - Mesh
FUNCTIONAL DESCRIPTION: BLGEOL-V1 software can generate hex-dominant meshes of geologic structures complying with different geometric constraints: surface topography (valleys, reliefs, rivers), geologic layers and underground workings. First, a reference 2D domain is obtained by projecting all the line constraints into a horizontal plane. Different size specifications are given for rivers, outcrop lines and workings. Using an adaptive methodology, the size variation is bounded by a specified threshold in order to obtain a high quality quad-dominant mesh. Secondly, a hex-dominant mesh of the geological medium is generated by a vertical extrusion, taking into account the surfaces found (interfaces between two layers, top or bottom faces of underground workings). The generation of volume elements follows a global order established on the whole set of surfaces to ensure the conformity of the resulting mesh.

- Participants: Houman Borouchaki and Patrick Laug
- Contact: Patrick Laug
- URL: http://pages.saclay.inria.fr/patrick.laug/logiciels/logiciels.html

4.6. BLMOL

KEYWORDS: Mesher - Molecular surface - Meshing - Mesh
SCIENTIFIC DESCRIPTION: An increasingly important part of quantum chemistry is devoted to molecular surfaces. To model such a surface, each constituting atom is idealized by a simple sphere. Surface mesh generation techniques are then used either for visualization or for simulation, where mesh quality has a strong influence on solution accuracy. First, a boundary representation (B-rep) of the surface is obtained, i.e. a set of patches and the topological relations between them. Second, an appropriate parameterization and a metric map are computed for each patch. Third, meshes of the parametric domains are generated with respect to an induced metric map, using a combined advancing-front generalized-Delaunay approach. Finally these meshes are mapped onto the entire surface. Several application examples illustrate various capabilities of our method.

FUNCTIONAL DESCRIPTION: BLMOL is a molecular surface mesher.

- Participants: Houman Borouchaki and Patrick Laug
- Contact: Patrick Laug
- URL: http://pages.saclay.inria.fr/patrick.laug/logiciels/logiciels.html

4.7. BLSURF

KEYWORDS: Automatic mesher - Meshing - Mesher - Mesh
FUNCTIONAL DESCRIPTION: An indirect method for meshing parametric surfaces conforming to a user-specifiable size map is used. First, from this size specification, a Riemannian metric is defined so that the desired mesh is one with unit length edges with respect to the related Riemannian space (the so-called

- Participants: Houman Borouchaki and Patrick Laug
- Partner: Université de Technologie de Troyes
- Contact: Patrick Laug
- URL: http://pages.saclay.inria.fr/patrick.laug/logiciels/logiciels.html

4.8. FEFLOA-REMESH

KEYWORDS: Scientific calculation - Anisotropic - Mesh adaptation
FUNCTIONAL DESCRIPTION: FEFLOA-REMESH is intended to generate adapted 2D, surface and volume meshes by using a unique cavity-based operator. The metric-aligned or metric-orthogonal approach is used to generate high quality surface and volume meshes independently of the anisotropy involved.

- Participants: Adrien Loseille and Frédéric Alauzet
- Contact: Adrien Loseille
- URL: https://www.rocq.inria.fr/gamma/Adrien.Loseille/index.php?page=softwares

4.9. GAMANIC 3D

KEYWORDS: Tetrahedral mesh - Delaunay - Anisotropic size and direction control - Automatic mesher
FUNCTIONAL DESCRIPTION: GAMANIC3D is a volume mesher governed by a (anisotropic) size and directional specification metric field.

- Participants: Adrien Loseille, Éric Saltel, Frédéric Alauzet, Frederic Hecht, Houman Borouchaki and Paul Louis George
- Contact: Paul Louis Georges
- URL: http://www.meshgems.com/volume-meshing.html

4.10. GAMHIC 3D

KEYWORDS: Tetrahedral mesh - Delaunay - Isotropic - Automatic mesher
FUNCTIONAL DESCRIPTION: GAMHIC3D is a volume mesher governed by a (isotropic) size specification metric field.

- Participants: Adrien Loseille, Éric Saltel, Frédéric Alauzet, Frederic Hecht, Houman Borouchaki and Paul Louis George
- Contact: Paul Louis George
- URL: http://www.meshgems.com/volume-meshing.html

4.11. GHS3D

KEYWORDS: Tetrahedral mesh - Delaunay - Automatic mesher

FUNCTIONAL DESCRIPTION: GHS3D is an automatic volume mesher

- Participants: Adrien Loseille, Éric Saltel, Frédéric Alauzet, Frederic Hecht, Houman Borouchaki and Paul Louis George
- Contact: Paul Louis George
- URL: http://www.meshgems.com/volume-meshing.html

4.12. HEXOTIC

KEYWORDS: 3D - Mesh generation - Meshing - Unstructured meshes - Octree/Quadtree - Multi-threading - GPGPU - GPU

FUNCTIONAL DESCRIPTION: Input: a triangulated surface mesh and an optional size map to control the size of inner elements.

Output: a fully hexahedral mesh (no hybrid elements), valid (no negative jacobian) and conformal (no dangling nodes) whose surface matches the input geometry.

The software is a simple command line that requires no knowledge on meshing. Its arguments are an input mesh and some optional parameters to control elements sizing, curvature and subdomains as well as some features like boundary layers generation.

- Participant: Loïc Maréchal
- Partner: Distene
- Contact: Loïc Maréchal
- URL: https://team.inria.fr/gamma3/project-presentation/gamma-software/hexotic/

4.13. Nimbus 3D

KEYWORDS: Surface reconstruction - Point cloud

FUNCTIONAL DESCRIPTION: Nimbus3D is a surface reconstruction method piece of software

- Participants: Houman Borouchaki and Paul Louis George
- Contact: Paul Louis George
- URL: http://www.meshgems.com/volume-meshing.html

4.14. VIZIR

Interactive visualization of hybrid, curbed and high-order mesh and solution

KEYWORD: Mesh

FUNCTIONAL DESCRIPTION: Vizir is a light, simple and interactive mesh visualization software, including : (i) A curved meshes visualizator: it handles high order elements and solutions, (ii) Hybrid elements mesh visualization (pyramids, prisms, hexahedra), (iii) Solutions visualization : clip planes, capping, iso-lines, iso-surfaces.

- Participants: Adrien Loseille and Rémi Feuillet
- Contact: Adrien Loseille
- Publication: Vizir: High-order mesh and solution visualization using OpenGL 4.0 graphic pipeline
- URL: http://vizir.inria.fr

4.15. Wolf

KEYWORD: Scientific calculation

FUNCTIONAL DESCRIPTION: Numerical solver for the Euler and compressible Navier-Stokes equations with turbulence modelling. ALE formulation for moving domains. Modules of interpolation, mesh optimisation and moving meshes. Wolf is written in C++, and may be later released as an opensource library. FELiScE was registered in July 2014 at the Agence pour la Protection des Programmes under the Inter Deposit Digital Number IDDN.FR.001.340034.000.S.P.2014.000.10000.

- Participants: Adrien Loseille and Frédéric Alauzet
- Contact: Frédéric Alauzet
- URL: http://pages.saclay.inria.fr/frederic.alauzet/code_eng.html#Wolf-Nsc

4.16. Wolf-Bloom

KEYWORD: Scientific calculation
FUNCTIONAL DESCRIPTION: Wolf-Bloom is a structured boundary layer mesh generator using a pushing approach. It start from an existing volume mesh and insert a structured boundary layer by pushing the volume mesh. The volume mesh deformation is solved with an elasticity analogy. Mesh-connectivity optimizations are performed to control volume mesh element quality.

- Participants: Adrien Loseille, David Marcum and Frédéric Alauzet
- Contact: Frédéric Alauzet
- URL: http://pages.saclay.inria.fr/frederic.alauzet/code_eng.html

4.17. Wolf-Elast

KEYWORD: Scientific calculation
FUNCTIONAL DESCRIPTION: Wolf-Elast is a linear elasticity solver using the P1 to P3 Finite-Element method. The Young and Poisson coefficient can be parametrized. The linear system is solved using the Conjugate Gradient method with the LUSGS preconditioner.

- Participants: Adrien Loseille and Frédéric Alauzet
- Contact: Frédéric Alauzet
- URL: http://pages.saclay.inria.fr/frederic.alauzet/code_eng.html

4.18. Wolf-Interpol

KEYWORD: Scientific calculation
FUNCTIONAL DESCRIPTION: Wolf-Interpol is a tool to transfer scalar, vector and tensor fields from one mesh to another one. Polynomial interpolation (from order 2 to 4) or conservative interpolation operators can be used. Wolf-Interpol also extract solutions along lines or surfaces.

- Participants: Adrien Loseille and Frédéric Alauzet
- Contact: Frédéric Alauzet
- URL: http://pages.saclay.inria.fr/frederic.alauzet/code_eng.html

4.19. Wolf-MovMsh

KEYWORD: Scientific calculation
FUNCTIONAL DESCRIPTION: Wolf-MovMsh is a moving mesh algorithm coupled with mesh-connectivity optimization. Mesh deformation is computed by means of a linear elasticity solver or a RBF interpolation. Smoothing and swapping mesh optimization are performed to maintain good mesh quality. It handles rigid bodies or deformable bodies, and also rigid or deformable regions of the domain. High-order meshes are also handled

- Participants: Adrien Loseille and Frédéric Alauzet
- Contact: Paul Louis George
- URL: http://pages.saclay.inria.fr/frederic.alauzet/code_eng.html

4.20. Wolf-Nsc

KEYWORD: Scientific calculation
FUNCTIONAL DESCRIPTION: Wolf-Nsc is numerical flow solver solving steady or unsteady turbulent compressible Euler and Navier-Stokes equations. The available turbulent models are the Spalart-Almaras and the Menter SST k-omega. A mixed finite volume - finite element numerical method is used for the discretization. Second order spatial accuracy is reached thanks to MUSCL type methods. Explicit or implicit time integration are available. It also resolved dual (adjoint) problem and compute error estimate for mesh adaptation.

- Participants: Adrien Loseille and Frédéric Alauzet
- Contact: Frédéric Alauzet
- URL: http://pages.saclay.inria.fr/frederic.alauzet/code_eng.html

4.21. Wolf-Spyder

KEYWORD: Scientific calculation
FUNCTIONAL DESCRIPTION: Wolf-Spyder is a metric-based high-order mesh quality optimizer using vertex smoothing and edge/face swapping.

- Participants: Adrien Loseille and Frédéric Alauzet
- Contact: Frédéric Alauzet
- URL: http://pages.saclay.inria.fr/frederic.alauzet/code_eng.html

5. New Results

5.1. The meshing bible

Participants: Paul-Louis George [The Boss], Houman Borouchaki, Frédéric Alauzet, Patrick Laug, Adrien Loseille, Loïc Maréchal.

Un projet important, initié en 2017, et amené à se poursuivre l'an prochain, consiste à écrire noir sur blanc un livre (en plusieurs volumes) et la motivation de ce travail est détaillée dans ce qui suit.

Pourquoi ce livre, pourquoi 2 volumes, pourquoi pas 3 volumes?

Notre dernier livre (généraliste) sur le maillage date de 2000 avec une mise à jour en 2008. Un collègue a commis un nouveau livre en 2015, très bien écrit mais assez classique dans son contenu, loin de préoccupations industrielles et (!) contenant quelques énormités (pas assez d'experience sur de vrais problèmes).

Ajoutons ma facilite (c'est P.L. G. qui parle) à écrire (bien ou mal, là n'est pas la question, il me suffit en effet de taper sur quelques touches d'un clavier), le désir de mon (premier) co-auteur de marquer le coup dans le domaine et la volonté (à leur corps defendant) des autres co-auteurs de participer à cette aventure. Le tout couplé avec les récents progrès dans le domaine (pensons aux éléments courbes et aux méthodes d'ordre élevé mais aussi à ce que peut être le HPC dans le domaine), tous les ingrédients sont là, on y va.

Le premier jet (un seul volume) se montre impossible à réaliser, il faudrait au minimum 800 pages, donc deux volumes a minima. Les deux volumes finis, ne reste il pas la place pour un troisième volume. Constatant avec effroi que nos étudiants (mais pas seulement) maitrisent bien force concepts mais sont incapables de voir, en pratique, comment les mettre en musique, le troisième volume est apparu comme une évidence (et on sera, au total, autour de 1000 pages).

A qui s'adresse ces volumes, bonne question. Ce n'est pas précisément de la littérature de gare mais nous nous sommes efforcé de prendre le malheureux lecteur par la main pour l'amener progressivement vers des concepts (très) avancés. Ainsi, le livre est très verbose et, en aucun cas, n'est un étalage savant de théorèmes et autres propositions, ce qui n'empêche pas de dire les choses. Par ailleurs, nous avons délibérément mis une part de subjectivité dans le propos pour suggérer (cela pouvant être contredit) que telle ou telle méthode n'avait pas notre faveur. A titre personnel, je pense que, bien que rares dans les livres, ces opinions ne peuvent qu'aider le lecteur a se former sa propre idée sur tel ou tel point.

Les livres sont publiés chez ISTE et écrits en français, eh oui, mais une traduction en anglais est avalable chez Wiley. La présence de la langue française dans la littérature scientifique me semble importante (et rejoint la politique de mon (notre) éditeur). Pour conclure, c'est plutôt satisfaisant de penser que ces livres (peut être destinés à faire référence sur le sujet) sont issus de l'Inria dans le neuf un.

5.2. Realistic modeling of fractured geologic media

Participants: Patrick Laug [correspondant], Géraldine Pichot.

This study started in 2016, in collaboration with the project-team Serena, aims to model, in a realistic and efficient manner, natural fractured media. These media are characterized by their diversity of structures and organizations. Numerous studies in the past decades have evidenced the existence of characteristic structures at multiple scales. At fracture scale, the aperture distribution is widely correlated and heterogeneous. At network scale, the topology is complex resulting from mutual mechanical interactions as well as from major stresses. Geometric modeling of fractured networks combines in a non-standard way a large number of 2D fractures interconnected in the 3D space. Intricate local configurations of fracture intersections require original methods of geometric modeling and mesh generation. Significant progress has been made during this year 2018, as we are now able to make geometric models and numerical simulations with more than 1 million fractures, 2 million intersections, and 18 million triangles, in about one hour on a laptop [7], [8], [19], [20], [21].

5.3. High order geometric modeling

Participants: Patrick Laug [correspondant], Houman Borouchaki.

In the area of geometric modeling, major challenges are linked to the efficient **visualization** of CAD surfaces and to the generation of **meshes** adapted to numerical simulation. In this context, the elaboration and implementation of a **discrete geometric model** provides a simple and universal representation model, without the need for CAD. A first study has been carried out for a model of degree 1 (one) defined by a "triangulation" composed of quadrilaterals and triangles. The advantage of this model of degree 1 lies in its geometric simplicity. However, in the case of complex surfaces, it may require a very large number of elements, and besides it is not sufficiently rich to give certain essential characteristics like geometric curvatures. The main goal of this project is to extend this discrete model of degree 1 to **higher degrees**. These studies are conducted by “MODIS”, an Associate Team comprising members of research teams at Inria, UTT (France) and Polytechnique Montreal (Canada) from 2017 to 2019. This year (2018) has been mostly devoted to the software implementation of all the theoretical bases obtained last year. In particular, chapters 6 and 11 of a recent book [22] give data structures where a local numbering is recursively defined for any order of the elements.

5.4. Rendu pixel-exact de solutions d'ordre élevé

Participants: Adrien Loseille [correspondant], Rémi Feuillet.

Avec le développement des méthodes d'ordre élevé, il apparaît également important de visualiser de manière fidèle à la fois le maillage et la solution associée. L'objet de cette thématique de recherche est de mettre à profit les fonctionnalités de programmation du pipeline graphique de la bibliothèque graphique OpenGL 4.0 afin de mettre en place des techniques de rendu de solutions d'ordre élevé quasiment exactes au pixel près ainsi que des techniques rapides de visualisation d'éléments d'ordre élevé. Les premiers résultats sont très satisfaisants avec un rendu de solutions d'ordre élevé allant de l'ordre 1 à 5 et ce exact au pixel près si ces dernières sont représentées sur des maillages de degré 1. Au niveau de la représentation des éléments d'ordre élevé, la visualisation est possible jusqu'à l'ordre 4 avec visualisation d'une solution dessus. Cette dernière sera d'autant plus représentée au pixel près que la représentation de la géométrie courbe (pour des degrés plus grands que 1) sera exacte. Ce travail a fait en 2018 l'objet de deux exposés à des conférences, dont une publication [13] dans les *proceedings*: *AIAA Scitech* et *WCCM*. Les prochaines étapes vont se concentrer sur la représentation de solutions et d'éléments d'ordre élevés à travers un plan de coupe, à la représentation des iso-surfaces et enfin à l'optimisation de ce code.

5.5. Génération de maillages d'ordre élevé

Participants: Frédéric Alauzet [correspondant], Adrien Loseille, Rémi Feuillet, David Marcum.

En calcul scientifique, l'utilisation de solveurs d'ordre élevé (supérieur à deux) se fait croissante. Or ces solveurs ne sont fonctionnels que lorsqu'ils sont couplés avec des maillages d'ordre élevé, nécessaires pour une représentation d'ordre élevé de la géométrie. L'idée de cette thématique de recherche est de s'intéresser à la génération et à l'*amélioration* par modification locale de tels maillages. Dans cette optique, un générateur de maillages courbe en partant de maillages droit a été développé. Ensuite, une étude importante a été consacrée à la généralisation au degré 2 des opérateurs locaux classiques d'optimisation de maillage, à savoir la bascule d'arête/face (*swap*) et le bougé de point (*smoothing*). La généralisation de tels outils a permis d'une part de rendre la génération de maillages courbe plus robuste et d'autre part rendu possibles au degré 2 les techniques de maillage mobile avec changement local de connectivité. Ce travail a fait l'objet en 2018 de deux exposés sans *proceedings* à *ECCM-ECFD* et *ICOSAHOM* et de 2 exposés avec *peer-reviewed proceedings* à *AIAA Aviation* et à l'*International Meshing Roundtable* [9], [10]. La suite de ce travail va être de générer des maillages de couche limites directement courbes (en utilisant la technique de maillage courbe mobile) puis de se consacrer à la généralisation de ce travail à des ordres plus grands. Il sera aussi apporté un soin particulier à la génération de maillages de surface courbe à partir de modèles de CAO.

5.6. Adaptation de maillages pour des écoulements visqueux en turbomachine et aéro-externe

Participants: Frédéric Alauzet, Loïc Frazza, Adrien Loseille [correspondant].

5.6.1. Calcul

Les principes d'une adaptation pour les écoulements Navier-Stokes turbulents ont été validés sur des calculs de turbomachine. Pour ce faire nous avons tout d'abord traité les particularités liées aux calculs en turbomachine:

- Les aubes présentent en général une périodicité par rotation, on ne simule donc qu'une période afin d'alléger les calculs. Il faut donc traiter cette périodicité de façon appropriée dans le code CFD et l'adaptation de maillage.
- Afin de prendre en compte la rotation des pales sans employer de maillages mobiles et simulations instationnaires on peut se placer dans le référentiel tournant de l'aube en corrigeant les équations.
- Les écoulements en turbomachine sont des écoulements clos, les conditions limites d'entrée et de sortie ont donc une influence très forte et peuvent de plus se trouver très près de la turbine afin de simuler la présence d'autres étages en amont ou en aval. Des conditions limites bien précises ont donc été développées afin de traiter correctement ces effets.

5.6.2. Adaptation

Pour l'adaptation de maillages deux particularités doivent ètre traitées ici, la périodicité du maillage et la couche limite turbulente. En toute dimemsion, la couche limite a donc été traitée par des techniques d'adaptation. Le maillage est adapté dans le volume en utilisant la Hessienne du Mach de l'écoulement comme senseur. La périodicité est traitée en utilisant un noyau non-manifold de logiciel de remaillage `Feflo.a`.

Ces développements ont été présentés dans plusieurs conférences internationales et sont détaillés dans la thèse de Loïc Frazza [1], et dans le cadre d'une collaboration avec Safran Tech.

5.7. Parallel mesh adaptation

Participants: Frédéric Alauzet, Adrien Loseille [correspondant].

We devise a strategy in order to generate large-size adapted anisotropic meshes $O(10^8 - 10^9)$ as required in many fields of application in scientific computing. We target moderate scale parallel computational resources as typically found in R&D units where the number of cores ranges in $O(10^2 - 10^3)$. Both distributed and shared memory architectures are handled. Our strategy is based on hierarchical domain splitting algorithm to remesh the partitions in parallel. Both the volume and the surface mesh are adapted simultaneously and the efficiency of the method is independent of the complexity of the geometry. The originality of the method relies on (i) a metric-based static load-balancing, (ii) dedicated hierarchical mesh partitioning techniques to (re)split the (complex) interfaces meshes, (iii) anisotropic Delaunay cavity to define the interface meshes, (iv) a fast, robust and generic sequential cavity-based mesh modification kernel, and (v) out-of-core storing of completing parts to reduce the memory footprint. We are able to generate (uniform, isotropic and anisotropic) meshes with more than 1 billion tetrahedra in less than 20 minutes on 120 cores.

5.8. Adaptive boundary layer mesh generation

Participants: Adrien Loseille [correspondant], Victorien Menier.

Si des méthodes traditionnelles de couches limites sont désormais matures, elles sont souvent incompatibles dans un contexte adaptatif où les tailles varient dans la couche limite. On a développé dans ce cadre une méthode d'adaptation de la couche limite à la fois dans la direction normale et dans le plan tangent basée sur un couplage entre un opérateur de cavité contraint et une adaptation de surface interne.

5.9. Améliorations des schémas pour les simulation RANS

Participants: Loïc Frazza, Frédéric Alauzet [correspondant].

Grâce à une implémentation adéquate de schémas numériques modernes, nous avons montré qu'il est possible de réaliser des simulations RANS sur des maillages tétraédriques non structurés. Nous avons ainsi pu réaliser des calculs sur des maillages adaptés pour différentes applications industrielles complexes. A cette fin, nous avons réalisé l'analyse mathématiques nécessaire au développement des senseurs d'erreurs turbulents efficaces et précis. Nous avons également été amenés à étendre la résolution des variables adjointes aux modèles RANS. En comparant les performances de cette stratégie d'adaptation, nous avons pu montrer la supériorité des résultats obtenus en comparaison des méthodes hessiennes et traditionnelles sur différentes applications. Tous ces développements ont été validés dans le solveur `Wolf`, présentés et publiés dans [12], [6], [15] et ont été développés dans le cadre des collaborations avec Boeing et Safran Tech.

5.10. Deterministic smoothing parallelization

Participants: Lucille-Marie Tenkès, Frédéric Alauzet [correspondant].

On élabore des solutions algorithmiques pour paralléliser un opérateur géométrique d'optimisation de maillage, le bougé de points : une solution non-dynamique et une solution dynamique.

Le but est de paralléliser les méthodes de bougé de points de manière déterministe. En effet, avec les algorithmes actuels, le résultat de l'optimisation dépend de l'ordre dans lequel les points sont traités, qui n'est pas prévisible en parallélisation multi-thread asynchrone. Cela devient problématique lorsqu'on insère cette étape dans un processus global. S'il survient une erreur, on ne pourra pas reproduire les cas invalides pour apporter des corrections. Une première idée a été de rendre les calculs d'optimisation déterministes en implémentant une méthode non dynamique, qui prend comme référence pour le calcul des positions optimales la configuration initiale et non la configuration en cours d'optimisation. Une relaxation est effectuée une fois toutes les positions optimales calculées, et elle est globale. Cela permet d'avoir un bougé de points vraiment indépendant de l'ordre dans lequel les points sont traités. Cette méthode est bien déterministe, mais l'algorithme est moins efficace et plus lent que les méthodes de bougé de points dynamiques. L'alternative dynamique consiste à agir directement sur la parallélisation, en regroupant les points de manière à ne pas laisser les calculs interférer. L'idée est de créer une partition des nœuds en mettant dans une même classe ceux qui ne sont pas reliés entre eux par une arête. Ainsi, si l'un d'entre eux bouge lors du processus d'optimisation, il n'impactera pas ceux de sa classe. Cela revient donc à colorier ces sommets, de sorte que deux nœuds reliés par une arête n'aient pas la même couleur. L'algorithme choisi effectue un coloriage de proche en proche (approche frontale). La couleur 1 est attribuée au premier sommet, les points de sa boule sont coloriés par élimination, et ainsi de suite. Ces algorithmes ont été testés sur des maillages 2D et 3D pour caractériser leurs performances. Il en ressort que la méthode dynamique est assez efficace et rapide pour être appliquée à des maillages 3D de grande taille. De plus, l'étape de partition a été elle-même parallélisée et optimisée.

5.11. Opérateurs d'optimisation de maillage alignés et maillages quad-dominants

Participants: Lucille-Marie Tenkès, Frédéric Alauzet [correspondant].

On se base ici sur les méthodes de maillage par alignement des éléments sur la métrique. C'est-à-dire que les éléments du maillage, en plus de présenter une taille adaptée à la solution, suivent une direction prescrite. Une des difficultés ressortant de cette méthode est l'optimisation par bougé de points, car cela rompt l'alignement. On veut donc ici mettre en place une technique de bougé de points permettant de corriger, sans en détruire la structure, un maillage généré par alignement sur la métrique. Les modifications apportées à l'optimiseur sont un opérateur de réduction d'arête, une procédure d'appariement en quadrilatères, et un opérateur de bougé de points utilisant ces quadrilatères.

5.12. Multi-physic mesh adaptation

Participants: Frédéric Alauzet [correspondant], Rémi Feuillet, Julien Vanharen.

A new strategy [18] for mesh adaptation dealing with Fluid-Structure Interaction (FSI) problems is presented using a partitioned approach. The Euler equations are solved by an edge-based Finite Volume solver whereas the linear elasticity equations are solved by the Finite Element Method using the Lagrange P1 elements. The coupling between both codes is realized by imposing boundary conditions. Small displacements of the structure are assumed and so the mesh is not deformed. The computation of a well-documented FSI test case is finally carried out to perform validation of this new strategy. The capability of treating three-dimensional complex cases is also demonstrated.

6. Bilateral Contracts and Grants with Industry

6.1. Bilateral Contracts with Industry

- Boeing
- Safran Tech

6.2. Bilateral Grants with Industry

- Projet RAPID DGA

7. Partnerships and Cooperations

7.1. National Initiatives

7.1.1. ANR

7.1.1.1. ANR IMPACTS 2018-2021

Ideal Mesh generation for modern solvers and comPuting ArchiteCTureS.

- Coordinateur : Adrien Loseille
- The rapid improvement of computer hardware and physical simulation capabilities has revolutionized science and engineering, placing computational simulation on an equal footing with theoretical analysis and physical experimentation. This rapidly increasing reliance on the predictive capabilities has created the need for rigorous control of numerical errors which strongly impact these predictions. A rigorous control of the numerical error can be only achieved through mesh adaptivity. In this context, the role of mesh adaptation is prominent, as the quality of the mesh, its refinement, and its alignment with the physics are major contributions to these numerical errors. The IMPACTS project aims at pushing the envelope in mesh adaptation in the context of large size, very high fidelity simulations by proposing a new adaptive mesh generation framework. This framework will be based on new theoretical developments on Riemannian metric-field and on innovative algorithmic developments coupling a unique cavity-operator with an advancing-point techniques in order to produce high quality hybrid, curved and adapted meshes.

7.2. International Initiatives

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

7.2.1.1. AM2NS

Title: Advanced Meshing Methods for Numerical Simulations

International Partner (Institution - Laboratory - Researcher):

Mississippi State University (United States) - Center for Advanced Vehicular Systems - Computational Fluid Dynamics Dept. (CAVS-CFD) - Marcum David

Start year: 2017

See also: http://pages.saclay.inria.fr/frederic.alauzet/AssociateTeam_AM2NS/AT_am2ns.html

The purpose of the AM2NS Associate Team is to mutualize the knowledge of all teams in order to develop the next generation of meshing methods and their parallelization to address the new challenges in numerical simulations for industrial problems. The Associate Team is composed of four partners: Inria, Mississippi State University, The Boeing Company and Massachusetts Institute of Technology.

7.2.1.2. MODIS

Title: High-order discrete geometric modeling

International Partner (Institution - Laboratory - Researcher):

Polytechnique Montréal (Canada) - Computer Science - Franoi̧s Guibault

Start year: 2017

See also: http://pages.saclay.inria.fr/patrick.laug/MODIS/MODIS.html

In the area of geometric modeling, major challenges are linked to the efficient visualization of CAD surfaces and to the generation of meshes adapted to numerical simulation. In this context, the conception of a discrete geometric model provides a simple and universal representation model, without the need for CAD. A first study has been carried out for the conception of a model of order 1 (one) defined by a 'triangulation' composed of quadrilaterals and triangles. The advantage of this model of order 1 lies in its geometric simplicity. However, in the case of complex surfaces, it may require a very large number of elements, and besides it is not sufficiently rich to give certain essential characteristics like geometric curvatures. The main goal of this project is to extend this discrete model of order 1 to higher orders.

8. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] L. FRAZZA.*3D anisotropic mesh adaptation for Reynolds Averaged Navier-Stokes simulations*, Sorbonne Université , UPMC, December 2018, https://hal.inria.fr/tel-01962318

Articles in International Peer-Reviewed Journal

[2] T. AMARI, A. CANOU, J.-J. ALY, F. DELYON, F. ALAUZET.*Magnetic cage and rope as the key for solar eruptions*, in "Nature", February 2018, vol. 554, n^{o} 7691, p. 211 - 215 [*DOI :* 10.1038/NATURE24671], https://hal-cea.archives-ouvertes.fr/cea-01872913

[3] S. CHAILLAT, S. P. GROTH, A. LOSEILLE.*Metric-based anisotropic mesh adaptation for 3D acoustic boundary element methods*, in "Journal of Computational Physics", November 2018, vol. 372, p. 473 - 499 [*DOI :* 10.1016/J.JCP.2018.06.048], https://hal.archives-ouvertes.fr/hal-01895636

[4] T. GROSGES, D. BARCHIESI.*Geometrical optimization of nanostrips for surface plasmon excitation: an analytical approach*, in "Optics Letters", 2018, vol. 43, n^o 1, p. 54-57 [*DOI :* 10.1364/OL.43.000054], https://hal.inria.fr/hal-01804193

[5] T. GROSGES, D. BARCHIESI.*Gold Nanoparticles as a Photothermal Agent in Cancer Therapy: The Thermal Ablation Characteristic Length*, in "Molecules", May 2018, vol. 23, n^o 6, 13 [*DOI :* 10.3390/MOLECULES23061316], https://hal.inria.fr/hal-01804012

[6] A. LOSEILLE, L. FRAZZA, F. ALAUZET.*Comparing anisotropic adaptive strategies on the 2nd AIAA sonic boom workshop geometry*, in "Journal of Aircraft", November 2018, p. 1-15, https://hal.inria.fr/hal-01962175

Invited Conferences

[7] G. PICHOT, P. LAUG, J. ERHEL, R. LE GOC, C. DARCEL, P. DAVY, J.-R. DE DREUZY.*Flow simulations in geology-based Discrete Fracture Networks*, in "2018 - Reactive Flows in Deformable, Complex Media", Oberwolfach, Germany, August 2018, p. 1-3, https://hal.inria.fr/hal-01900605

[8] G. PICHOT, P. LAUG, R. LE GOC, C. DARCEL, P. DAVY, J.-R. DE DREUZY.*Computation of flow properties of large scale fractured media*, in "InterPore 2018 - 10th Annual Meeting and Jubilee", New Orleans, United States, June 2018, https://hal.inria.fr/hal-01900599

International Conferences with Proceedings

[9] R. FEUILLET, A. LOSEILLE, F. ALAUZET.*P2 mesh optimization operators*, in "27th International Meshing Roundtable", Albuquerque, United States, October 2018, https://hal.inria.fr/hal-01962132

[10] R. FEUILLET, A. LOSEILLE, D. MARCUM, F. ALAUZET.*Connectivity-change moving mesh methods for high-order meshes: Toward closed advancing-layer high-order boundary layer mesh generation*, in "2018 Fluid Dynamics Conference, AIAA AVIATION Forum", Atlanta, United States, June 2018 [*DOI :* 10.2514/6.2018-4167], https://hal.inria.fr/hal-01962129

[11] L. FRAZZA, A. LOSEILLE, F. ALAUZET, A. DERVIEUX.*Nonlinear corrector for RANS equations*, in "2018 Fluid Dynamics Conference, AIAA AVIATION Forum", Atlanta, France, June 2018, https://hal.inria.fr/hal-01962171

[12] L. FRAZZA, A. LOSEILLE, F. ALAUZET.*Mesh adaptation strategies using wall functions and low-Reynolds models*, in "2018 Fluid Dynamics Conference, AIAA AVIATION Forum", Atlanta, United States, June 2018, https://hal.inria.fr/hal-01962178

[13] A. LOSEILLE, R. FEUILLET.*Vizir: High-order mesh and solution visualization using OpenGL 4.0 graphic pipeline*, in "2018 - AIAA Aerospace Sciences Meeting, AIAA SciTech Forum", kissimmee, United States, January 2018, p. 1-13 [*DOI :* 10.2514/6.2018-1174], https://hal.inria.fr/hal-01686714

[14] A. LOSEILLE.*Recent Improvements on Cavity-Based Operators for RANS Mesh Adaptation*, in "2018 AIAA Aerospace Sciences Meeting", Kissimmee, United States, January 2018, https://hal.inria.fr/hal-01962190

[15] T. A. MICHAL, F. ALAUZET, A. LOSEILLE, L. FRAZZA, D. MARCUM, D. KAMENETSKIY.*Comparing Anisotropic Error Estimates for the Onera M6 Wing RANS Simulations*, in "2018 AIAA Aerospace Sciences Meeting", Kissimmee, United States, January 2018, https://hal.inria.fr/hal-01962254

[16] E. MONTREUIL, W. GHEDHAIFI, V. CHMIELARSKI, V. FRANÇOIS, F. GAND, A. LOSEILLE.*Numerical Simulation of contrail formation on the Common Research Model wing/body/engine configuration*, in "AIAA Aviation and Aeronautics Forum and Exposition 2018", Atlanta, United States, June 2018, https://hal.archives-ouvertes.fr/hal-01961143

[17] M. A. PARK, N. BARRAL, D. IBANEZ, D. KAMENETSKIY, J. J. KRAKOS, T. A. MICHAL, A. LOSEILLE.*Unstructured Grid Adaptation and Solver Technology for Turbulent Flows*, in "2018 AIAA Aerospace Sciences Meeting", Kissimmee, United States, January 2018, https://hal.inria.fr/hal-01962247

[18] J. VANHAREN, R. FEUILLET, F. ALAUZET.*Mesh adaptation for fluid-structure interaction problems*, in "2018 Fluid Dynamics Conference, AIAA AVIATION Forum", Atlanta, United States, June 2018 [*DOI :* 10.2514/6.2018-3244], https://hal.inria.fr/hal-01962204

Conferences without Proceedings

[19] P. LAUG, G. PICHOT, R. LE GOC, C. DARCEL, P. DAVY.*Automatic meshing of Discrete Fracture Networks*, in "Computational Methods in Water Resources XXII (CMWR 2018)", Saint-Malo, France, June 2018, https://hal.inria.fr/hal-01896927

[20] P. LAUG, G. PICHOT.*Simulations in large tridimensional Discrete Fracture Networks (DFN): I. Geometric modeling and mesh generation*, in "MASCOT 2018 - 15th IMACS/ISGG meeting on applied scientific computing and tools", Rome, Italy, October 2018, p. 1-2, https://hal.inria.fr/hal-01896881

[21] G. PICHOT, P. LAUG, J. ERHEL, R. LE GOC, C. DARCEL, P. DAVY, J.-R. DE DREUZY.*Simulations in large tridimensional Discrete Fracture Networks (DFN): II. Flow simulations*, in "MASCOT 2018 -15th IMACS/ISGG meeting on applied scientific computing and tools", Rome, Italy, October 2018, https://hal.inria.fr/hal-01896900

Scientific Books (or Scientific Book chapters)

[22] P.-L. GEORGE, H. BOROUCHAKI, F. ALAUZET, P. LAUG, A. LOSEILLE, L. MARÉCHAL.*Maillage, modélisation géométrique et simulation numérique 2*, Iste Editions, 2018, https://hal.inria.fr/hal-01962201

Research Reports

[23] H. GUILLARD, J. LAKHLILI, A. LOSEILLE, A. LOYER, B. NKONGA, A. RATNANI, A. ELARIF.*Tokamesh : A software for mesh generation in Tokamaks*, CASTOR, December 2018, n^o RR-9230, https://hal.inria.fr/hal-01948060

Other Publications

[24] A. BELME, F. ALAUZET, A. DERVIEUX.*An a priori anisotropic Goal-Oriented Error Estimate for Viscous Compressible Flow and Application to Mesh Adaptation*, November 2018, working paper or preprint, https://hal.inria.fr/hal-01927113

[25] A. DERVIEUX, E. GAUCI, L. FRAZZA, A. BELME, A. CARABIAS, A. LOSEILLE, F. ALAUZET.*Mesh-Anpassung für k-genaue Approximationen in CFD*, November 2018, working paper or preprint, https://hal.inria.fr/hal-01927145

[26] E. GAUCI, A. BELME, A. CARABIAS, A. LOSEILLE, F. ALAUZET, A. DERVIEUX.*A priori error-based mesh adaptation in CFD*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01928249

[27] J. V. LANGENHOVE, D. LUCOR, F. ALAUZET, A. BELME.*Goal-oriented error control of stochastic system approximations using metric-based anisotropic adaptations*, February 2018, https://arxiv.org/abs/1805.00370 - working paper or preprint, https://hal.inria.fr/hal-01703054

References in notes

[28] P.-G. CIARLET.*Basic Error Estimates for Elliptic Problems*, Ciarlet, P. G. and Lions, J. L., North Holland, 1991, vol. II

[29] P.-J. FREY, P.-L. GEORGE.*Maillages. Applications aux éléments finis*, Hermès Science Publications, Paris, 1999

Project-Team GRACE

Geometry, arithmetic, algorithms, codes and encryption

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

IN PARTNERSHIP WITH:
CNRS

Ecole Polytechnique

RESEARCH CENTER
Saclay - Île-de-France

THEME
Algorithmics, Computer Algebra and Cryptology

Table of contents

Project-Team GRACE

Creation of the Team: 2012 January 01, updated into Project-Team: 2013 July 01

Keywords:

Computer Science and Digital Science:

A4.2. - Correcting codes
A4.3. - Cryptography
A4.3.1. - Public key cryptography
A4.3.3. - Cryptographic protocols
A4.3.4. - Quantum Cryptography
A4.8. - Privacy-enhancing technologies
A8.1. - Discrete mathematics, combinatorics
A8.4. - Computer Algebra
A8.5. - Number theory

Other Research Topics and Application Domains:

B9.5.1. - Computer science
B9.5.2. - Mathematics
B9.10. - Privacy

1. Team, Visitors, External Collaborators

Research Scientists

Daniel Augot [Team leader, Inria, Senior Researcher, HDR]
Alain Couvreur [Inria, Researcher]
Benjamin Smith [Inria, Researcher]
Matthieu Rambaud [Inria, Researcher, en détachement depuis le corps Mines-Télécom, until Sep 2018]

Faculty Members

Luca de Feo [Université de Versailles Saint-Quentin-en-Yvelines, Associate Professor, en délégation depuis l'Université de Versailles Saint-Quentin, until Aug 2018]
Françoise Levy-Dit-Vehel [École Nationale Supérieure de Techniques Avancées, Professor, HDR]
François Morain [Ecole polytechnique, Professor, HDR]

External Collaborators

Luca de Feo [Université de Versailles Saint-Quentin-en-Yvelines, from Sep 2018]
Elise Barelli [Univ de Versailles Saint-Quentin-en-Yvelines, from Sep 2018]

Technical Staff

Nicholas Coxon [Inria, until nov 2018]

PhD Students

Elise Barelli [Inria, until Aug 2018]
Lucas Benmouffok [Institut de recherche technologique System X, from Oct 2018]
Hanna-Mae Bisserier [Institut de recherche technologique System X]
Sarah Bordage [Ecole polytechnique, from Oct 2018]
Mathilde Chenu de La Morinerie [Ecole polytechnique, from Oct 2018]
Hussein Khazaie [Inria, until May 2018]
Julien Lavauzelle [Ecole polytechnique until Sep 2018, Inria from Oct 2018]
Isabella Panaccione [Inria, from Oct 2018]

Post-Doctoral Fellow
William George [Ecole polytechnique, until Jan 2018]

Administrative Assistants
Jessica Gameiro [Inria, until Apr 2018]
Maria Agustina Ronco [Inria, from May 2018]

2. Overall Objectives

2.1. Scientific foundations

GRACE has two broad application domains—cryptography and coding theory—linked by a common foundation in algorithmic number theory and the geometry of algebraic curves. In our research, which combines theoretical work with practical software development, we use algebraic curves to *create better cryptosystems*, to *provide better security assessments* for cryptographic key sizes, and to *build the best error-correcting codes*.

Coding and cryptography deal (in different ways) with securing communication systems for high-level applications. In our research, the two domains are linked by the computational issues related to algebraic curves (over various fields) and arithmetic rings. These fundamental number-theoretic algorithms, at the crossroads of a rich area of mathematics and computer science, have already proven their relevance in public key cryptography, with industrial successes including the RSA cryptosystem and elliptic curve cryptography. It is less well-known that the same branches of mathematics can be used to build very good codes for error correction. While coding theory has traditionally had an electrical engineering flavour, recent developments in computer science have shed new light on coding theory, leading to new applications more central to computer science.

3. Research Program

3.1. Algorithmic Number Theory

Algorithmic Number Theory is concerned with replacing special cases with general algorithms to solve problems in number theory. In the Grace project, it appears in three main threads:

- fundamental algorithms for integers and polynomials (including primality and factorization);
- algorithms for finite fields (including discrete logarithms); and
- algorithms for algebraic curves.

Clearly, we use computer algebra in many ways. Research in cryptology has motivated a renewed interest in Algorithmic Number Theory in recent decades—but the fundamental problems still exist *per se*. Indeed, while algorithmic number theory application in cryptanalysis is epitomized by applying factorization to breaking RSA public key, many other problems, are relevant to various area of computer science. Roughly speaking, the problems of the cryptological world are of bounded size, whereas Algorithmic Number Theory is also concerned with asymptotic results.

3.2. Arithmetic Geometry: Curves and their Jacobians

Theme: Arithmetic Geometry: Curves and their Jacobians *Arithmetic Geometry* is the meeting point of algebraic geometry and number theory: that is, the study of geometric objects defined over arithmetic number systems (such as the integers and finite fields). The fundamental objects for our applications in both coding theory and cryptology are curves and their Jacobians over finite fields.

An algebraic *plane curve*$\mathcal{X}$ over a field $\mathbf{K}$ is defined by an equation

$$\mathcal{X} : F_{\mathcal{X}}(x, y) = 0 \quad \text{where } F_{\mathcal{X}} \in \mathbf{K}[x, y].$$

(Not every curve is planar—we may have more variables, and more defining equations—but from an algorithmic point of view, we can always reduce to the plane setting.) The *genus*$g_{\mathcal{X}}$ of $\mathcal{X}$ is a non-negative integer classifying the essential geometric complexity of $\mathcal{X}$; it depends on the degree of $F_{\mathcal{X}}$ and on the number of singularities of $\mathcal{X}$. The curve $\mathcal{X}$ is associated in a functorial way with an algebraic group $J_{\mathcal{X}}$, called the *Jacobian* of $\mathcal{X}$. The group $J_{\mathcal{X}}$ has a geometric structure: its elements correspond to points on a $g_{\mathcal{X}}$-dimensional projective algebraic group variety. Typically, we do not compute with the equations defining this projective variety: there are too many of them, in too many variables, for this to be convenient. Instead, we use fast algorithms based on the representation in terms of classes of formal sums of points on $\mathcal{X}$.

The simplest curves with nontrivial Jacobians are curves of genus 1, known as *elliptic curves*; they are typically defined by equations of the form $y^2 = x^3 + Ax + B$. Elliptic curves are particularly important given their central role in public-key cryptography over the past two decades. Curves of higher genus are important in both cryptography and coding theory.

3.3. Curve-Based cryptology

Theme: Curve-Based Cryptology

Jacobians of curves are excellent candidates for cryptographic groups when constructing efficient instances of public-key cryptosystems. Diffie–Hellman key exchange is an instructive example.

Suppose Alice and Bob want to establish a secure communication channel. Essentially, this means establishing a common secret *key*, which they will then use for encryption and decryption. Some decades ago, they would have exchanged this key in person, or through some trusted intermediary; in the modern, networked world, this is typically impossible, and in any case completely unscalable. Alice and Bob may be anonymous parties who want to do e-business, for example, in which case they cannot securely meet, and they have no way to be sure of each other's identities. Diffie–Hellman key exchange solves this problem. First, Alice and Bob publicly agree on a cryptographic group G with a generator P (of order N); then Alice secretly chooses an integer a from $[1..N]$, and sends aP to Bob. In the meantime, Bob secretly chooses an integer b from $[1..N]$, and sends bP to Alice. Alice then computes $a(bP)$, while Bob computes $b(aP)$; both have now computed abP, which becomes their shared secret key. The security of this key depends on the difficulty of computing abP given P, aP, and bP; this is the Computational Diffie–Hellman Problem (CDHP). In practice, the CDHP corresponds to the Discrete Logarithm Problem (DLP), which is to determine a given P and aP.

This simple protocol has been in use, with only minor modifications, since the 1970s. The challenge is to create examples of groups G with a relatively compact representation and an efficiently computable group law, and such that the DLP in G is hard (ideally approaching the exponential difficulty of the DLP in an abstract group). The Pohlig–Hellman reduction shows that the DLP in G is essentially only as hard as the DLP in its largest prime-order subgroup. We therefore look for compact and efficient groups of prime order.

The classic example of a group suitable for the Diffie–Hellman protocol is the multiplicative group of a finite field $\mathbf{F}_q$. There are two problems that render its usage somewhat less than ideal. First, it has too much structure: we have a subexponential Index Calculus attack on the DLP in this group, so while it is very hard, the DLP falls a long way short of the exponential difficulty of the DLP in an abstract group. Second, there is only one such group for each q: its subgroup treillis depends only on the factorization of $q - 1$, and requiring $q - 1$ to have a large prime factor eliminates many convenient choices of q.

This is where Jacobians of algebraic curves come into their own. First, elliptic curves and Jacobians of genus 2 curves do not have a subexponential index calculus algorithm: in particular, from the point of view of the DLP, a generic elliptic curve is currently *as strong as* a generic group of the same size. Second, they provide some diversity: we have many degrees of freedom in choosing curves over a fixed $\mathbf{F}_q$, with a consequent diversity of possible cryptographic group orders. Furthermore, an attack which leaves one curve vulnerable may not necessarily apply to other curves. Third, viewing a Jacobian as a geometric object rather than a pure group allows us to take advantage of a number of special features of Jacobians. These features include efficiently computable pairings, geometric transformations for optimised group laws, and the availability of efficiently computable non-integer endomorphisms for accelerated encryption and decryption.

3.4. Algebraic Coding Theory

Theme: Coding theory

Coding Theory studies originated with the idea of using redundancy in messages to protect against noise and errors. The last decade of the 20th century has seen the success of so-called iterative decoding methods, which enable us to get very close to the Shannon capacity. The capacity of a given channel is the best achievable transmission rate for reliable transmission. The consensus in the community is that this capacity is more easily reached with these iterative and probabilistic methods than with algebraic codes (such as Reed–Solomon codes).

However, algebraic coding is useful in settings other than the Shannon context. Indeed, the Shannon setting is a random case setting, and promises only a vanishing error probability. In contrast, the algebraic Hamming approach is a worst case approach: under combinatorial restrictions on the noise, the noise can be adversarial, with strictly zero errors.

These considerations are renewed by the topic of list decoding after the breakthrough of Guruswami and Sudan at the end of the nineties. List decoding relaxes the uniqueness requirement of decoding, allowing a small list of candidates to be returned instead of a single codeword. List decoding can reach a capacity close to the Shannon capacity, with zero failure, with small lists, in the adversarial case. The method of Guruswami and Sudan enabled list decoding of most of the main algebraic codes: Reed–Solomon codes and Algebraic–Geometry (AG) codes and new related constructions “capacity-achieving list decodable codes”. These results open the way to applications again adversarial channels, which correspond to worst case settings in the classical computer science language.

Another avenue of our studies is AG codes over various geometric objects. Although Reed–Solomon codes are the best possible codes for a given alphabet, they are very limited in their length, which cannot exceed the size of the alphabet. AG codes circumvent this limitation, using the theory of algebraic curves over finite fields to construct long codes over a fixed alphabet. The striking result of Tsfasman–Vladut–Zink showed that codes better than random codes can be built this way, for medium to large alphabets. Disregarding the asymptotic aspects and considering only finite length, AG codes can be used either for longer codes with the same alphabet, or for codes with the same length with a smaller alphabet (and thus faster underlying arithmetic).

From a broader point of view, wherever Reed–Solomon codes are used, we can substitute AG codes with some benefits: either beating random constructions, or beating Reed–Solomon codes which are of bounded length for a given alphabet.

Another area of Algebraic Coding Theory with which we are more recently concerned is the one of Locally Decodable Codes. After having been first theoretically introduced, those codes now begin to find practical applications, most notably in cloud-based remote storage systems.

4. Application Domains

4.1. Internet of Things

The *Internet of Things* (IoT) is the network and application space formed by the millions of small, connected devices that are increasingly present in our daily lives, and by the servers, clouds, and apps that they communicate with. This includes not only consumer devices such as smartphones, household devices, and wearable technology, but also an increasinly large proportion of our fundamental civic infrastructure (as is reflected by the increasing attention given to *Smart Cities*).

The IoT is therefore a massive, pervasive, and highly heterogeneous distributed computing system; a system that is mostly unprotected and insecure. Many of the devices are simply too small and underpowered to run the conventional cryptosystems that are standard for internet communications: even a minimalist TLS stack will often overwhelm the resources available on some small platforms. These limitations include small memory size, limited battery power, and low computational capacity. Not only are these devices harder to defend, but they are also much easier to attack: for example, these devices are generally extremely physically accessible (they must be, to fulfil their purpose), but this makes them extremely vulnerable to side-channel attacks.

Nevertheless, strong cryptography is essential to the future of IoT, precisely because these systems are so pervasive in our everyday lives, both individually (in our homes) and collectively (in our cities, industries, and urban infrastructure). We need strong cryptography to protect the personal and industrial data that these devices collect, process, and transmit; but we also need strong cryptography to ensure that devices and services can identify and authenticate themselves and each other with confidence. It is not enough to simply put secure systems in place; we must also develop reliable software update mechanisms, tailored to the needs and challenges of the IoT space.

While these technical challenges have been met, to some extent, for symmetric cryptosystems (which means that we have reasonable means of encrypting data and ensuring its integrity), they pose a massive problem for implementers of asymmetric cryptosystems (including key exchange, signatures, identification, and authentication). Efficient asymmetric cryptosystems have long been a research focus for GRACE, and our expertise in elliptic curve cryptosystems is of particular relevance for IoT, since these cryptosystems typically require the fewest memory and bandwidth resources.

Looking towards the future, the massive contemporary research effort in postquantum cryptosystems has so far mostly yielded systems even less-suited to IoT than conventional asymmetric systems are. Nevertheless, there is some hope that postquantum security can be brought to some IoT devices, and we are hopeful that GRACE's strength in isogeny-based cryptography will have an impact here.

4.2. Cloud storage

The team is concerned with several aspect of reliability and security of cloud storage, obtained mainly with tools from coding theory. On the privacy side, we build protocols for so-called Private Information Retrieval which enable a user to query a remote database for an entry, while not revealing his query. For instance, a user could query a service for stock quotes without revealing with company he is interested in. On the availability side, we study protocols for proofs of retrievability, which enable a user to get assurance that a huge file is still available on a remote server, with a low bandwith protocol which does not require to download the whole file. For instance, in a peer-to-peer distributed storage system, where nodes could be rewarded for storing data, they can be audited with proof of retrievability protocols to make sure they indeed hold the data.

We investigate these problems with algebraic coding theory, mainly codes with locality (locally decodable codes, locally recoverable codes, and so on).

4.3. Blockchains

The huge interest shown by companies for blockchains and cryptocurrencies have attracted the attention of mainstream industries for new, advanced uses of cryptographic, beyond confidentiality, integrity and authentication. In particular, zero-knowledge proofs, computation with encrypted data, etc, are now revealing their potential in the blockchain context. Team Grace is investigating two topics in these areas: secure multiparty computation and so-called "STARKS".

Secure multiparty computation enables several participants to compute a common function of data they each secretly own, without each participant revealing his data to the other participants. This area has seen great progress in recent years, and the cryptogaphic protocols are now mature enough for practical use. This topic is new to project-team Grace, and we will investigate it in the context of blockchains, through the lenses of use for private "smart contracts". A PhD student has been hired since October, funded by IRT System-X.

(ZK-)STARKS stands for "(Zero-Knowledge) Scalable Transparent ARguments of Knowledge", which can be zero knowledge or not. These techniques enable to have short probabilistic proof of correctness of program execution, which can be quicly checked by a verifier, without requiring the verifier to redo the computation again. This topic is close to the problem of computational integrity, and its theoretical foundations originate back to the 90's, which saw the formulation and proof of the celebrated PCP theorem. A protocol family equivalent of STARKS, "SNARKS", are well established, performant and promoted by the zerocash protocol for anomymous cryptocurrency (and also available in Ethereum), and STARKS are seen as a future replacement for SNARKS, overcoming the SNARKS problem of trusted setup. At the core of STARKS lie algebraic codes, mainly basic Reed-Solomon codes, and we will investigate replacement for the Reed-Solomon codes, to allow more performant (shorter) STARKS.

5. New Software and Platforms

5.1. ACTIS

Algorithmic Coding Theory in Sage
FUNCTIONAL DESCRIPTION: The aim of this project is to vastly improve the state of the error correcting library in Sage. The existing library does not present a good and usable API, and the provided algorithms are very basic, irrelevant, and outdated. We thus have two directions for improvement: renewing the APIs to make them actually usable by researchers, and incorporating efficient programs for decoding, like J. Nielsen's CodingLib, which contains many new algorithms.

- Partner: Technical University Denmark
- Contact: Daniel Augot

5.2. DECODING

KEYWORD: Algebraic decoding
FUNCTIONAL DESCRIPTION: Decoding is a standalone C library. Its primary goal is to implement Guruswami–Sudan list decoding-related algorithms, as efficiently as possible. Its secondary goal is to give an efficient tool for the implementation of decoding algorithms (not necessarily list decoding algorithms) and their benchmarking.

- Participant: Guillaume Quintin
- Contact: Daniel Augot

5.3. Fast Compact Diffie-Hellman

KEYWORD: Cryptography
FUNCTIONAL DESCRIPTION: A competitive, high-speed, open implementation of the Diffie–Hellman protocol, targeting the 128-bit security level on Intel platforms. This download contains Magma files that demonstrate how to compute scalar multiplications on the x-line of an elliptic curve using endomorphisms. This accompanies the EuroCrypt 2014 paper by Costello, Hisil and Smith, the full version of which can be found here: http://eprint.iacr.org/2013/692 . The corresponding SUPERCOP-compatible crypto_dh application can be downloaded from http://hhisil.yasar.edu.tr/files/hisil20140318compact.tar.gz .

- Participant: Benjamin Smith
- Contact: Benjamin Smith
- URL: http://research.microsoft.com/en-us/downloads/ef32422a-af38-4c83-a033-a7aafbc1db55/

5.4. CADO-NFS

Crible Algébrique: Distribution, Optimisation - Number Field Sieve
KEYWORDS: Cryptography - Number theory
FUNCTIONAL DESCRIPTION: CADO-NFS is a complete implementation in C/C++ of the Number Field Sieve (NFS) algorithm for factoring integers and computing discrete logarithms in finite fields. It consists in various programs corresponding to all the phases of the algorithm, and a general script that runs them, possibly in parallel over a network of computers.
NEWS OF THE YEAR: The main program for relation collection now supports composite "special-q", and also parallelizes better. The memory footprint of the central step of linear algebra has been reduced, and the parallelism of this step has been improved.

- Participants: Pierrick Gaudry, Emmanuel Thomé and Paul Zimmermann
- Contact: Emmanuel Thomé
- URL: http://cado-nfs.gforge.inria.fr/

6. New Results

6.1. Fast transforms over fields of characteristic 2

Participant: Nicholas Coxon.

With the aim of reaching fast, linear time, algorithms for encoding multiplicity codes, which have good local properties, N. Coxon had to develop subalgorithms for dealing with the Hermite interpolation [13], which in turn relies on computer algebra for fast transforms over fields of characteritic two [14]. Locally decodable codes are used for private information retrieval, where a database can be privately queried by a user, in such a way that the user does not reveal his query. Using codes with locality for private information retrieval, the database is first encoded, then queried using the local property of the code. Since the databases in question can be large, only linear time algorithms can be used. Our results achieve linear-time complexity, and even with a non agressively optimized implementation, can encode as much as 10^9 bits in thirty seconds on a laptop.

6.2. Private information retrieval

Participants: Daniel Augot, Nicholas Coxon, Julien Lavauzelle, Françoise Levy-Dit-Vehel.

J. Lavauzelle continued his study on private information retrieval (PIR) protocols. First, he completed the construction of PIR protocols from transversal designs [8], initiated in 2017. Compared to existing protocols, the main benefit of the construction is to feature an optimal computation complexity for the servers. Sublinear communication complexity and negligeable storage overhead can also be achieved for some particular instances.

Second, in a joint work with R. Tajeddine, R. Freij-Hollanti and C. Hollanti from the University of Aalto (Finland), J. Lavauzelle considered the setting in which the database is encoded with an optimal regenerating code [16]. Quantitatively, their construction of PIR protocols improves upon a recent work of Dorkson and Ng, for every non-trivial set of parameters.

6.3. Locally correctable codes

Participant: Julien Lavauzelle.

In 2013, Guo, Kopparty and Sudan built a new family of locally correctable codes from lifting, achieving an arbitrarily high information rate for sublinear locality. J. Lavauzelle proposed an analogue of this construction in projective spaces [7]. The parameters of this construction are similar to the original work of Guo *et al.* Intertwined relations between the two families of codes were proven thanks to a careful analysis of their monomial bases. The practicality of the construction was also established through an implementation and a study of information sets and automorphisms of the code.

6.4. Cryptanalysis in code based cryptography

Participant: Alain Couvreur.

Following NIST call for post quantum cryptography, A. Couvreur and E. Barelli designed a key recovery attack against a McEliece–like encryption scheme called DAGS [9].

In addition, in collaboration with Matthieu Lequesne and Jean-Pierre Tillich (Inria Paris, SECRET team), A. Couvreur designed an attack against another proposal called RLCE (Random Linear Code Encryption) [12].

6.5. Commutative isogeny-based cryptography

Participants: Luca de Feo, Benjamin Smith.

Despite the many advances in post-quantum cryptography in recent years, efficient drop-in replacements for the classic Diffie–Hellman key exchange algorithm have proven elusive. L. De Feo, J. Kieffer, and B. Smith laid the algorithmic groundwork for *commutative isogeny-based key exchange* in [10]; this work became the basis of the exciting new CSIDH proposal [19].

6.6. Factoring oracles

Participants: François Morain, Benjamin Smith.

Integer factoring is an old topic, and the situation is as follows: in the classical world, we think integer factoring is hard and the algorithms we have are quite powerful though of subexponential complexity and factoring numbers with several hundred bits; whereas in the quantum world, it is assumed to be easy (i.e., there exists a quantum polynomial time algorithm) but never experienced and the record is something like a few bits. F. Morain, helped by B. Smith and G. Renault (ANSSI) studied the theoretical problem of factoring integers given access to classical oracles, like the Euler totient function. They were able to give some interesting classes of numbers that could tackled, see [17].

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

7.1.1. Nokia

Participants: Daniel Augot, Nicholas Coxon, Françoise Levy-Dit-Vehel.

Phase 2 has been finished, while a new phase, phase 3, has been negociated between Inria and Nokia. Grace finished his work on fast algorithms for polynomials over fields of small caracteristic, wth application to coding theory, multiplicity codes and private information retrieval. The new phase will fund a project on rank-metric codes for security and privacy in cloud storage (in collaboration with Gilles Zémor, Uni. Bordeaux).

8. Partnerships and Cooperations

8.1. Regional Initiatives

Participants: Daniel Augot, Matthieu Rambaud.

A “research initiative” “BART” (Blockchain advanced research and technologies) has been launched with three partners: Inria, Institut Mines-Télécom, and System-X. This is funded by *Institut de recherche* System-X, located in Paris-Saclay area, whose objective is to connect industry and academia. A new PhD has been started, with L. Benmouffok, hired in October 2018, whose topic is the use of secure multiparty computation in blockchains.

8.2. National Initiatives

8.2.1. ANR

Participants: Daniel Augot, Alain Couvreur, Matthieu Rambaud.

MANTA (accepted July 2015, starting March 2016): "Curves, surfaces, codes and cryptography". This project deals with applications of coding theory error correcting codes to in cryptography, multi-party computation, and complexity theory, using advanced topics in algebraic geometry and number theory. The kickoff was a one week-retreat in Dordogne (20 participants), and we had another four day meeting in Saclay in November 17. See http://anr-manta.inria.fr/.

8.3. European Initiatives

8.3.1. SPARTA

- Program: H2020
- Project acronym: SPARTA
- Project title: SPARTA
- Duration: three years
- Coordinator: CEA
- Other partners: IMT, Inria, ANSSI
- Abstract: Propose, test, validate and exploit the possible organizational, technological and operational setup of a cybersecurity competence network; Produce a roadmap that include targets to be achieved by the end of the project, as well as priorities to be addressed in the future by the Cybersecurity Competence Network; Serve to align research, education and certification; Build on and align existing roadmap efforts.
  Participant: Benjamin Smith.

8.3.2. PQCRYPTO

Title: Post-quantum cryptography for long-term security

Programm: H2020

Duration: March 2015 - March 2018

Coordinator: TECHNISCHE UNIVERSITEIT EINDHOVEN

Partners:

Academia Sinica (Taiwan)

Bundesdruckerei (Germany)

Danmarks Tekniske Universitet (Denmark)

Katholieke Universiteit Leuven (Belgium)

Nxp Semiconductors Belgium Nv (Belgium)

Ruhr-Universitaet Bochum (Germany)

Stichting Katholieke Universiteit (Netherlands)

Coding Theory and Cryptology group, Technische Universiteit Eindhoven (Netherlands)

Technische Universitaet Darmstadt (Germany)

University of Haifa (Israel)

Inria contact: Nicolas Sendrier

Online security depends on a very few underlying cryptographic algorithms. Essentially all applications today are based on RSA or on the discrete-logarithm problem in finite fields or on elliptic curves. Cryptographers optimize parameter choices and implementation details for these systems and build protocols on top of these systems; cryptanalysts fine-tune attacks and establish exact security levels for these systems.

These systems are all broken as soon as large quantum computers are built. Long-term confidential documents such as patient health-care records and state secrets have to guarantee security for many years, but information encrypted today using RSA or elliptic curves and stored until quantum computers are available will then be as easy to decipher.

PQCRYPTO will allow users to switch to post-quantum cryptography: PQCRYPTO will design a portfolio of high-security post-quantum public-key systems, and will improve the speed of these systems, with reference implementations.

Our team is engaged in WP3.3 “advanced applications for the cloud”. We envision to focus essentially on secure multiparty computation, essentially the information theoretically secure constructions, who are naturally secure against a quantum computer invoked on classical queries. We will study whether these protocols still resist quantum queries. This work sub package started March 2015, ended in March 2018.
Participants: Daniel Augot, Matthieu Rambaud.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Selection

9.1.1.1. Member of the Conference Program Committees

- D. Augot was in the program committee of FAB 2018, Foundations and Applications of Blockchain, Los Angeles.
- D. Augot was in the program committee of WTSC 2018, Workshop on Trusted Smart Contracts, Curaçao.
- D. Augot was in the program committee of WAIFI 2018, Workshop on the Arithmetic of Finite Fields, Bergen, Norway.
- D. Augot was in the program committee of BCT 2018, International Workshop on Cryptocurrencies and Blockchain Technology, in conjunction with ESORICS 2018, Barcelona.
- A. Couvreur was in the program committee of the *Journées codes et cryptographie (C2) 2018*.

9.1.1.2. Reviewer

- D. Augot: ISIT 2018 (International Symposium on Information Theory)
- B. Smith: ANTS 2018, Indocrypt 2018, PKC 2019

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards

- F. Morain is member of the editorial board of the *Applicable Algebra in Engineering, Communication and Computing*, Springer.
- With Thomas Johansson, Marine Minier, Faina Soloveva, Victor Zinonviev, D. Augot is guest editor for a special issue of *Designs, Codes and Cryptography*, devoted to WCC2017, Workshop on Coding and Cryptography, St Petersburg, Russia.

9.1.2.2. Reviewer - Reviewing Activities

- A. Couvreur: Designs, Codes and Cryptography, Asiacrypt 2018, IEEE Transactions on information theory, Advances in Mathematics of communication, etc...
- J. Lavauzelle: Designs, Codes and Cryptography (special issue WCC 2017)
- B. Smith: Designs, Codes, and Cryptography, Finite Fields and their Applications, Journal of the London Mathematical Society, Mathematics of Computation,

9.1.3. Invited Talks

- D. Augot was an invited speaker of the Munich Workshop on Coding and Cryptography (MWCC) 2018
- D. Augot was an invited speaker at ACA 2018, Application of Computer Algebra, Santiago de Compostela
- D. Augot was invited at Dasgsthul Seminar 18511, Algebraic Coding Theory for Networks, Storage, and Security, and gave here a talk.
- B. Smith was an invited speaker at the *International Workshop on the Arithmetic of Finite Fields (WAIFI 2018)* (Bergen, Norway).
- B. Smith was an invited speaker at the *Journées Codage et Cryptographie 2018* (Aussois, France).

9.1.4. Industrial Show

- F. Levy-dit-Vehel demoed our Private Information Retrieval protocol at "FIC", International Security Forum, Lille, January 2018.

9.1.5. Leadership within the Scientific Community

- D. Augot is member of the scientific committee of the C2-CCA seminar, held three or four times a year, with a France wide audience, and which is the seminar of "groupe de travail" C2 "codage et cryptographie" of the GDR IM "groupement de recherche informatique mathématique".
- D. Augot is leading the scientific committee of the blocksem seminar of Plateau de Saclay.

9.1.6. Scientific Expertise

- A. Couvreur was evaluator for research grants attribution by university of Crete.

9.1.7. Research Administration

- F. Morain is vice-head of the Département d'informatique of Ecole Polytechnique; in charge of years 1 and 2 for Computer Science courses.
- F. Morain is member of the Board of Master Parisien de Recherche en Informatique (MPRI).
- A. Couvreur is member of Inria Saclay *Commission Scientifique*.
- D. Augot was member of the jury for two Inria Grenoble Rhône-Alpes positions
- D. Augot was member of the jury for a position at Institut Mines-Télécom.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence :

- F. Morain, Lectures for INF311: "Introduction à l'informatique", 15h (equiv TD), 1st year (L3), École polytechnique. Coordinator of this module (350 students).
- J. Lavauzelle, *Éléments de programmation* (1I002), 13.5h, L1, Université Pierre et Marie Curie, France
- A. Couvreur, INF411 *Introduction à la programmation et à l'algorithmique*, 40h, L3, École Polytechnique, France

- B. Smith, CSE101 *Introduction to Computer Programming*, 36h, L1, École polytechnique, France

Master :

- F. Morain is the scientific leader of the Graduate Degree *Cybersecurity: Threats and Defense* of École Polytechnique.
- A. Couvreur, *Coding theory and application to cryptography*, 20h, M2, MPRI (Université Paris VII, ENS Paris, ENS Cachan, École Polytechnique), France
- F. Morain and A. Couvreur, INF558, *Introduction to cryptology*, 36h, M1, École Polytechnique.
- B. Smith, INF568 *Advanced Cryptography*, 36h, M1, École polytechnique
- B. Smith and F. Morain, *Algorithmes arithmétiques pour la cryptologie*, 20h, M2, MPRI (Université Paris VII, ENS Paris, ENS Cachan, École Polytechnique), France
- F. Levy-dit-Vehel, discrete maths, 21h, M1, ENSTA.
- F. Levy-dit-Vehel, cryptography, 24h, M2, ENSTA.

Doctorat :

- A. Couvreur, *Introduction to code based cryptography*, 6 hours. Spring school *Post Scryptum*

9.2.2. Supervision

- PhD : J. Lavauzelle, *Codes à propriétés locales : constructions et applications à des protocoles cryptographiques*, Université Paris Saclay.
- PhD : E. Barelli, *Étude de la sécurité de certaines clés compactes pour le schéma de McEliece utilisant des codes géométriques*, Université Paris Saclay.

9.2.3. Juries

- D. Augot, A. Couvreur, and F. Levy-dit-Vehelwere in the jury of J. Lavauzelle's PhD defense, le 30 novembre 2018, à Palaiseau: *Codes à propriétés locales : constructions et applications à des protocoles cryptographiques*
- D. Augot and A. Couvreur were in the jury of E. Barelli's PhD defense, le 10 décembre 2018 à Palaiseau: *Étude de la sécurité de certaines clés compactes pour le schéma de McEliece utilisant des codes géométriques*
- D. Augot was in in the committee of
  - Victor Cauchois, le jeudi 13 Décembre 2018 à Rennes: *Couches de diffusion linéaires à partir de matrices MDS*
  - Sviat Covanov, le 5 juin 2018 à Nancy: *Multiplication algorithms: algebraic complexity and fast asymptotic methods*
  - Jonathan Detchart, le 5 décembre 2018, à Toulouse: *Optimisation de codes correcteurs d'effacements par application de transformées polynomiales*

9.3. Popularization

9.3.1. Internal or external Inria responsibilities

- D. Augot is member of the "comité de pilotage" the "BART" (Blockchain advanced research and technologies) research initiative, with Institut Mines Télécom and System-X.

9.3.2. Interventions

- D. Augot was interviewed on blockchains by three representatives of the French National Assembly.

- D. Augot was interviewed by “France Stratégie”, an institution attached to the Prime Minister to support forward thinking of the French government.
- F. Levy-dit-Vehel demoed our Private Information Retrieval protocol with partitionned locally decodable codes

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] E. BARELLI.*Étude de la sécurité de certaines clés compactes pour le schéma de McEliece utilisant des codes géométriques*, Université Paris-Saclay, December 2018, https://pastel.archives-ouvertes.fr/tel-01982502

[2] J. LAVAUZELLE.*Codes with locality : constructions and applications to cryptographic protocols*, Université Paris-Saclay, November 2018, https://pastel.archives-ouvertes.fr/tel-01951078

Articles in International Peer-Reviewed Journal

[3] B. AUDOUX, A. COUVREUR.*On tensor products of CSS Codes*, in "Annales de l'Institut Henri Poincaré (D) Combinatorics, Physics and their Interactions", 2018, https://arxiv.org/abs/1512.07081 , https://hal.archives-ouvertes.fr/hal-01248760

[4] D. AUGOT, P. LOIDREAU, G. ROBERT.*Generalized Gabidulin codes over fields of any characteristic*, in "Designs, Codes and Cryptography", 2018, vol. 86, n^o 8, p. 1807-1848, https://arxiv.org/abs/1703.09125 [*DOI :* 10.1007/S10623-017-0425-6], https://hal.archives-ouvertes.fr/hal-01503212

[5] C. BACHOC, A. COUVREUR, G. ZÉMOR.*Towards a function field version of Freiman's Theorem*, in "Algebraic Combinatorics", 2018, vol. 1, n^o 4, p. 501-521, https://arxiv.org/abs/1709.00087 [*DOI :* 10.5802/ALCO.19], https://hal.archives-ouvertes.fr/hal-01584034

[6] E. BARELLI, P. BEELEN, M. DATTA, V. NEIGER, J. ROSENKILDE.*Two-Point Codes for the Generalized GK Curve*, in "IEEE Transactions on Information Theory", 2018 [*DOI :* 10.1109/TIT.2017.2763165], https://hal.archives-ouvertes.fr/hal-01535513

[7] J. LAVAUZELLE.*Lifted projective Reed–Solomon codes*, in "Designs, Codes and Cryptography", 2018, https://arxiv.org/abs/1809.00931 [*DOI :* 10.1007/S10623-018-0552-8], https://hal.archives-ouvertes.fr/hal-01901147

[8] J. LAVAUZELLE.*Private Information Retrieval from Transversal Designs*, in "IEEE Transactions on Information Theory", 2018, 1, https://arxiv.org/abs/1709.07952 [*DOI :* 10.1109/TIT.2018.2861747], https://hal.archives-ouvertes.fr/hal-01901014

International Conferences with Proceedings

[9] E. BARELLI, A. COUVREUR.*An efficient structural attack on NIST submission DAGS*, in "ASIACRYPT 2018", Brisbane, Australia, Advances in Cryptology – ASIACRYPT 2018, December 2018, vol. 11272, https://arxiv.org/abs/1805.05429 [*DOI :* 10.1007/978-3-030-03326-2_4], https://hal.archives-ouvertes.fr/hal-01796338

[10] L. DE FEO, J. KIEFFER, B. SMITH.*Towards practical key exchange from ordinary isogeny graphs*, in "ASIACRYPT 2018", Brisbane, Australia, December 2018, https://arxiv.org/abs/1809.07543 , https://hal.inria.fr/hal-01872817

Other Publications

[11] A. COUVREUR, M. LEQUESNE, J.-P. TILLICH.*Recovering short secret keys of RLCE encryption scheme in polynomial time*, October 2018, JC2 2018 - Journées Codage et Cryptographie, https://hal.inria.fr/hal-01959617

[12] A. COUVREUR, M. LEQUESNE, J.-P. TILLICH.*Recovering short secret keys of RLCE in polynomial time*, May 2018, https://arxiv.org/abs/1805.11489 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01803440

[13] N. COXON.*Fast Hermite interpolation and evaluation over finite fields of characteristic two*, July 2018, https://arxiv.org/abs/1807.00645 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01827583

[14] N. COXON.*Fast transforms over finite fields of characteristic two*, July 2018, https://arxiv.org/abs/1807.07785 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01845238

[15] S. GALBRAITH, L. PANNY, B. SMITH, F. VERCAUTEREN.*Quantum Equivalence of the DLP and CDHP for Group Actions*, December 2018, https://arxiv.org/abs/1812.09116 - working paper or preprint, https://hal.inria.fr/hal-01963660

[16] J. LAVAUZELLE, R. TAJEDDINE, R. FREIJ-HOLLANTI, C. HOLLANTI.*Private Information Retrieval Schemes with Regenerating Codes*, November 2018, https://arxiv.org/abs/1811.02898 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01951956

[17] F. MORAIN, G. RENAULT, B. SMITH.*Deterministic factoring with oracles*, February 2018, https://arxiv.org/abs/1802.08444 - working paper or preprint, https://hal.inria.fr/hal-01715832

[18] B. SMITH.*Pre- and post-quantum Diffie-Hellman from groups, actions, and isogenies*, September 2018, https://arxiv.org/abs/1809.04803 - working paper or preprint, https://hal.inria.fr/hal-01872825

References in notes

[19] W. CASTRYCK, T. LANGE, C. MARTINDALE, L. PANNY, J. RENES.*CSIDH: An Efficient Post-Quantum Commutative Group Action*, in "Advances in Cryptology - ASIACRYPT 2018 - 24th International Conference on the Theory and Application of Cryptology and Information Security, Brisbane, QLD, Australia, December 2-6, 2018, Proceedings, Part III", T. PEYRIN, S. D. GALBRAITH (editors), Lecture Notes in Computer Science, Springer, 2018, vol. 11274, p. 395–427, https://doi.org/10.1007/978-3-030-03332-3_15

Project-Team ILDA

Interacting with Large Data

IN COLLABORATION WITH: Laboratoire de recherche en informatique (LRI)

IN PARTNERSHIP WITH:
CNRS

Université Paris-Sud (Paris 11)

RESEARCH CENTER
Saclay - Île-de-France

THEME
Interaction and visualization

Table of contents

Project-Team ILDA

Creation of the Team: 2015 January 01, updated into Project-Team: 2016 December 01

Keywords:

Computer Science and Digital Science:

A3.1.7. - Open data
A3.1.10. - Heterogeneous data
A3.2.4. - Semantic Web
A3.2.5. - Ontologies
A3.2.6. - Linked data
A5.1. - Human-Computer Interaction
A5.2. - Data visualization
A5.5.4. - Animation
A5.6.1. - Virtual reality
A5.6.2. - Augmented reality

Other Research Topics and Application Domains:

B9.5.3. - Physics
B9.5.6. - Data science
B9.6.7. - Geography
B9.7.2. - Open data
B9.11. - Risk management

1. Team, Visitors, External Collaborators

Research Scientists

Emmanuel Pietriga [Team leader, Inria, Senior Researcher, HDR]
Caroline Appert [CNRS, Senior Researcher, HDR]
Olivier Chapuis [CNRS, Researcher]

Faculty Member

Anastasia Bezerianos [Univ Paris-Sud, Associate Professor]

External Collaborator

Bruno Fruchard [PhD student, Telecom ParisTech]

Technical Staff

Hande Gözükan [Inria, SED, until Sep 2018]
Léo Colombaro [Inria, from Sep 2018]
Adhitya Kamakshidasan [Inria, from Jun 2018]
Dylan Lebout [Inria]

PhD Students

Eugénie Brasier [Univ Paris-Sud, from Oct 2018 (intern from Apr 2018 to Sep 2018)]
Marie Destandau [Inria]
Anna Gogolou [Inria]
Raphaël James [Univ Paris-Sud, from Oct 2018 (intern from Apr 2018 to Sep 2018)]
Hugo Romat [TKM (CIFRE)]
Tong Xue [Inria, from Oct 2018]

Post-Doctoral Fellow

Vanessa Peña Araya [Inria, from Dec 2018]

Administrative Assistant

Alexandra Merlin [Inria]

2. Overall Objectives

2.1. Overall Objectives

In an increasing number of domains, computer users are faced with large datasets, that are often interlinked and organized according to elaborate structures thanks to new data models such as those that are arising with the development of, *e.g.*, the Web of Data. Rather than seeing the inherent complexity of those data models as a hindrance, we aim at leveraging it to design new interactive systems that can better assist users in their data understanding and processing tasks.

These "Data-centric Interactive Systems" aim at providing users with the right information at the right time, presenting it in the most meaningful manner, and letting users efficiently manipulate, edit and share these data with others. This entails minimizing the effort required to retrieve and relate data from relevant sources; displaying data using visual presentation techniques that match the data's characteristics and the users' tasks; and providing users with means of interacting with the data that effectively support their train of thought.

Our approach is based on the idea of bringing the fields of Web data management [35] and Human-computer interaction [59], [81] closer together, based on the strong belief that they have the potential to cross-fertilize one another. User interface design is essential to the management and understanding of large, interlinked datasets. Interlinked datasets enriched with even a small amount of semantics have the potential to help create interfaces that let users analyze and manipulate data in a more efficient manner by providing them with, *e.g.*, more relevant query results and giving them efficient means to navigate and relate those results. Our ultimate, long-term goal is to design interactive systems that make it as straightforward to manipulate large webs of data as spreadsheets do for tabular data.

3. Research Program

3.1. Introduction

Our ability to acquire or generate, store, process, interlink and query data has increased spectacularly over the last few years. The corresponding advances are commonly grouped under the umbrella of so called *Big Data*. Even if the latter has become a buzzword, these advances are real, and they are having a profound impact in domains as varied as scientific research, commerce, social media, industrial processes or e-government. Yet, looking ahead, emerging technologies related to what we now call the *Web of Data* (a.k.a the Semantic Web) have the potential to create an even larger revolution in data-driven activities, by making information accessible to machines as semistructured data [34] that eventually becomes actionable knowledge. Indeed, novel Web data models considerably ease the interlinking of semi-structured data originating from multiple independent sources. They make it possible to associate machine-processable semantics with the data. This in turn means that heterogeneous systems can exchange data, infer new data using reasoning engines, and that software agents can cross data sources, resolving ambiguities and conflicts between them [79]. Datasets are becoming very rich and very large. They are gradually being made even larger and more heterogeneous, but also much more useful, by interlinking them, as exemplified by the Linked Data initiative [54].

These advances raise research questions and technological challenges that span numerous fields of computer science research: databases, communication networks, security and trust, data mining, as well as human-computer interaction. Our research is based on the conviction that interactive systems play a central role in many data-driven activity domains. Indeed, no matter how elaborate the data acquisition, processing and storage pipelines are, data eventually get processed or consumed one way or another by users. The latter are faced with large, increasingly interlinked heterogeneous datasets (see, *e.g.*, Figure 1) that are organized according to complex structures, resulting in overwhelming amounts of both raw data and structured information. Users thus require effective tools to make sense of their data and manipulate them.

Figure 1. Linking Open Data cloud diagram from 2007 to 2017 – http://lod-cloud.net

We approach this problem from the perspective of the Human-Computer Interaction (HCI) field of research, whose goal is to study how humans interact with computers and inspire novel hardware and software designs aimed at optimizing properties such as efficiency, ease of use and learnability, in single-user or cooperative work contexts. More formally, HCI is about designing systems that lower the barrier between users' cognitive model of what they want to accomplish, and computers' understanding of this model. HCI is about the design, implementation and evaluation of computing systems that humans interact with [59], [81]. It is a highly multidisciplinary field, with experts from computer science, cognitive psychology, design, engineering, ethnography, human factors and sociology.

In this broad context, ILDA aims at designing interactive systems that display [43], [66], [88] the data and let users interact with them, aiming to help users better *navigate* and *comprehend* large webs of data represented visually, as well as *relate* and *manipulate* them.

Our research agenda consists of the three complementary axes detailed in the following subsections. Designing systems that consider interaction in close conjunction with data semantics is pivotal to all three axes. Those semantics will help drive navigation in, and manipulation of, the data, so as to optimize the communication bandwidth between users and data.

3.2. Semantics-driven Data Manipulation

Participants: Emmanuel Pietriga, Caroline Appert, Anastasia Bezerianos, Marie Destandau, Hugo Romat, Tong Xue, Léo Colombaro, Hande Gözükan.

The Web of Data has been maturing for the last fifteen years and is starting to gain adoption across numerous application domains (Figure 1). Now that most foundational building blocks are in place, from knowledge representation, inference mechanisms and query languages [55], all the way up to the expression of data presentation knowledge [74] and to mechanisms like look-up services [87] or spreading activation [50], we need to pay significant attention to how human beings are going to interact with this new Web, if it is to *"reach its full potential"* [51].

Most efforts in terms of user interface design and development for the Web of data have essentially focused on tools for software developers or subject-matter experts who create ontologies and populate them [61], [48]. Tools more oriented towards end-users are starting to appear [40], [42], [56], [57], [60], [69], including the so-called *linked data browsers* [54]. However, those browsers are in most cases based on quite conventional point-and-click hypertext interfaces that present data to users in a very page-centric, web-of-documents manner that is ill-suited to navigating in, and manipulating, webs of data.

To be successful, interaction paradigms that let users navigate and manipulate data on the Web have to be tailored to the radically different way of browsing information enabled by it, where users directly interact with the data rather than with monolithic documents. The general research question addressed in this part of our research program is how to design novel interaction techniques that help users manipulate their data more efficiently. By data manipulation, we mean all low-level tasks related to manually creating new content, modifying and cleaning existing content, merging data from different sources, establishing connections between datasets, categorizing data, and eventually sharing the end results with other users; tasks that are currently considered quite tedious because of the sheer complexity of the concepts, data models and syntax, and the interplay between all of them.

Our approach is based on the conviction that there is a strong potential for cross-fertilization, as mentioned earlier: on the one hand, user interface design is essential to the management and understanding of webs of data; on the other hand, interlinked datasets enriched with even a small amount of semantics can help create more powerful user interfaces, that provide users with the right information at the right time.

We envision systems that focus on the data themselves, exploiting the underlying *semantics and structure* in the background rather than exposing them – which is what current user interfaces for the Web of Data often do. We envision interactive systems in which the semantics and structure are not exposed directly to users, but serve as input to the system to generate interactive representations that convey information relevant to the task at hand and best afford the possible manipulation actions.

Relevant publications by team members this year: [22], [29], [28].

3.3. Generalized Multi-scale Navigation

Participants: Caroline Appert, Anastasia Bezerianos, Olivier Chapuis, Emmanuel Pietriga, Vanessa Peña Araya, Marie Destandau, Anna Gogolou, Hugo Romat, Adhitya Kamakshidasan, Dylan Lebout.

The foundational question addressed here is what to display when, where and how, so as to provide effective support to users in their data understanding and manipulation tasks. ILDA targets contexts in which workers have to interact with complementary views on the same data, or with views on different-but-related datasets, possibly at different levels of abstraction. Being able to combine or switch between representations of the data at different levels of detail and merge data from multiple sources in a single representation is central to many scenarios. This is especially true in both of the application domains we consider: mission-critical systems (e.g., natural disaster crisis management) and the exploratory analysis of scientific data (e.g., correlate theories and heterogeneous observational data for an analysis of a given celestial body in Astrophysics).

A significant part of our research over the last ten years has focused on multi-scale interfaces. We designed and evaluated novel interaction techniques, but also worked actively on the development of open-source UI toolkits for multi-scale interfaces (http://zvtm.sf.net). These interfaces let users navigate large but relatively homogeneous datasets at different levels of detail, on both workstations [77], [37], [73], [72], [71], [38], [76], [36], [78] and wall-sized displays [68], [63], [75], [67], [39], [45], [44]. This part of the ILDA research program is about extending multi-scale navigation in two directions: 1. Enabling the representation of multiple, spatially-registered but widely varying, multi-scale data layers in Geographical Information Systems (GIS); 2. Generalizing the multi-scale navigation paradigm to interconnected, heterogeneous datasets as found on the Web of Data.

The first research problem has been mainly investigated in collaboration with IGN in the context of ANR project MapMuxing (Section 8.1.1), which stands for *multi-dimensional map multiplexing*. Project MapMuxing aims at going beyond the traditional pan & zoom and overview+detail interface schemes, and at designing and evaluating novel cartographic visualizations that rely on high-quality generalization, *i.e.*, the simplification of geographic data to make it legible at a given map scale [84], [85], and symbol specification. Beyond project MapMuxing, we are also investigating multi-scale multiplexing techniques for geo-localized data in the specific context of ultra-high-resolution wall-sized displays, where the combination of a very high pixel density and large physical surface (Figure 3) enable us to explore designs that involve collaborative interaction and physical navigation in front of the workspace. This is work done in cooperation with team Massive Data at Inria Chile.

The second research problem is about the extension of multi-scale navigation to interconnected, heterogeneous datasets. Generalization has a rather straightforward definition in the specific domain of geographical information systems, where data items are geographical entities that naturally aggregate as scale increases. But it is unclear how generalization could work for representations of the more heterogeneous webs of data that we consider in the first axis of our research program. Those data form complex networks of resources with multiple and quite varied relationships between them, that cannot rely on a single, unified type of representation (a role played by maps in GIS applications).

Addressing the limits of current generalization processes is a longer-term, more exploratory endeavor. Here again, the machine-processable semantics and structure of the data give us an opportunity to rethink how users navigate interconnected heterogeneous datasets. Using these additional data, we investigate ways to generalize the multi-scale navigation paradigm to datasets whose layout and spatial relationships can be much richer and much more diverse than what can be encoded with static linear hierarchies as typically found today in interfaces for browsing maps or large imagery. Our goal is thus to design and develop highly dynamic and versatile multi-scale information spaces for heterogeneous data whose structure and semantics are not known in advance, but discovered incrementally.

Relevant publications by team members this year: [4], [25], [24], [15], [13], [12], [31], [11], [14].

3.4. Novel Forms of Input for Groups and Individuals

Participants: Caroline Appert, Anastasia Bezerianos, Olivier Chapuis, Emmanuel Pietriga, Eugénie Brasier, Bruno Fruchard, Raphaël James.

Analyzing and manipulating large datasets can involve multiple users working together in a coordinated manner in multi-display environments: workstations, handheld devices, wall-sized displays [39]. Those users work towards a common goal, navigating and manipulating data displayed on various hardware surfaces in a coordinated manner. Group awareness [53], [33] is central in these situations, as users, who may or may not be co-located in the same room, can have an optimal individual behavior only if they have a clear picture of what their collaborators have done and are currently doing in the global context. We work on the design and implementation of interactive systems that improve group awareness in co-located situations [62], making individual users able to figure out what other users are doing without breaking the flow of their own actions.

In addition, users need a rich interaction vocabulary to handle large, structured datasets in a flexible and powerful way, regardless of the context of work. Input devices such as mice and trackpads provide a limited number of input actions, thus requiring users to switch between modes to perform different types of data manipulation and navigation actions. The action semantics of these input devices are also often too much dependent on the display output. For instance, a mouse movement and click can only be interpreted according to the graphical controller (widget) above which it is moved. We focus on designing powerful input techniques based upon technologies such as tactile surfaces (supported by UI toolkits developed in-house), 3D motion tracking systems, or custom-built controllers [65] *to complement (rather than replace) traditional input devices* such as keyboards, that remain the best method so far for text entry, and indirect input devices such as mice or trackpads for pixel-precise pointing actions.

The input vocabularies we investigate enable users to navigate and manipulate large and structured datasets in environments that involve multiple users and displays that vary in their size, position and orientation [39], [52], each having their own characteristics and affordances: wall displays [68], [90], workstations, tabletops [70], [47], tablets [7], [86], smartphones [89], [46], [82], [83], and combinations thereof [2], [10], [67], [39].

We aim at designing rich interaction vocabularies that go far beyond what current touch interfaces offer, which rarely exceeds five gestures such as simple slides and pinches. Designing larger gesture vocabularies requires identifying discriminating dimensions (e.g., the presence or absence of anchor points and the distinction between internal and external frames of reference [7]) in order to structure a space of gestures that interface designers can use as a dictionary for choosing a coherent set of controls. These dimensions should be few and simple, so as to provide users with gestures that are easy to memorize and execute. Beyond gesture complexity, the scalability of vocabularies also depends on our ability to design robust gesture recognizers that will allow users to fluidly chain simple gestures that make it possible to interlace navigation and manipulation actions.

We also study how to further extend input vocabularies by combining touch [7], [89], [70] and mid-air gestures [68] with physical objects [58], [80], [65] and classical input devices such as keyboards to enable users to input commands to the system or to involve other users in their workflow (request for help, delegation, communication of personal findings, etc.) [41], [64]. Gestures and objects encode a lot of information in their shape, dynamics and direction, that can be directly interpreted in relation with the user, independently from the display output. Physical objects can also greatly improve coordination among actors for, e.g., handling priorities or assigning specific roles.

Relevant publications by team members this year: [25], [21], [17], [23], [20], [19], [27].

4. Application Domains

4.1. Mission-critical systems

Mission-critical contexts of use include emergency response & management, and critical infrastructure operations, such as public transportation systems, communications and power distribution networks, or the operations of large scientific instruments such as particle accelerators and astronomical observatories. Central to these contexts of work is the notion of situation awareness [33], i.e., how workers perceive and understand elements of the environment with respect to time and space, such as maps and geolocated data feeds from the field, and how they form mental models that help them predict future states of those elements. One of the main challenges is how to best assist subject-matter experts in constructing correct mental models and making informed decisions, often under time pressure. This can be achieved by providing them with, or helping them efficiently identify and correlate, relevant and timely information extracted from large amounts of raw data, taking into account the often cooperative nature of their work and the need for task coordination. With this application area, our goal is to investigate novel ways of interacting with computing systems that improve collaborative data analysis capabilities and decision support assistance in a mission-critical, often time-constrained, work context.

Relevant publications by team members this year: [25], [31], [16], [21], [23], [28], [13].

4.2. Exploratory analysis of scientific data

Many scientific disciplines are increasingly data-driven, including astronomy, molecular biology, particle physics, or neuroanatomy. While making the right decision under time pressure is often less of critical issue when analyzing scientific data, at least not on the same temporal scale as truly time-critical systems, scientists are still faced with large-to-huge amounts of data. No matter their origin (experiments, remote observations, large-scale simulations), these data are difficult to understand and analyze in depth because of their sheer size and complexity. Challenges include how to help scientists freely-yet-efficiently explore their data, keep a trace of the multiple data processing paths they considered to verify their hypotheses and make it easy to backtrack, and how to relate observations made on different parts of the data and insights gained at different moments during the exploration process. With this application area, our goal is to investigate how data-centric interactive systems can improve collaborative scientific data exploration, where users' goals are more open-ended, and where roles, collaboration and coordination patterns [53] differ from those observed in mission-critical contexts of work.

Relevant publications by team members last year: [15], [23], [30].

5. New Software and Platforms

5.1. Smarties

FUNCTIONAL DESCRIPTION: The Smarties system provides an easy way to add mobile interactive support to collaborative applications for wall displays.

It consists of (i) a mobile interface that runs on mobile devices for input, (ii) a communication protocol between the mobiles and the wall application, and (iii) libraries that implement the protocol and handle synchronization, locking and input conflicts. The library presents the input as an event loop with callback functions and handles all communication between mobiles and wall application. Developpers can customize the mobile interface from the wall application without modifying the mobile interface code.

On each mobile we find a set of cursor controllers associated with keyboards, widgets and clipboards. These controllers (pucks) can be shared by multiple collaborating users. They can control simple cursors on the wall application, or specific content (objects or groups of them). The developper can decide the types of widgets associated to pucks from the wall application side.

- Contact: Olivier Chapuis
- URL: http://smarties.lri.fr/

5.2. ZVTM

Zoomable Visual Transformation Machine
KEYWORDS: Big data - Visualization - Data visualization - Information visualization - Graph visualization
FUNCTIONAL DESCRIPTION: ZVTM is a toolkit enabling the implementation of multi-scale interfaces for interactively navigating in large datasets displayed as 2D graphics.

ZVTM is used for browsing large databases in multiple domains: geographical information systems, control rooms of complex facilitites, astronomy, power distribution systems.

The toolkit also enables the development of applications running on ultra-high-resolution wall-sized displays.

- Participants: Caroline Appert, Olivier Chapuis and Emmanuel Pietriga
- Contact: Emmanuel Pietriga
- Publications: Rapid Development of User Interfaces on Cluster-Driven Wall Displays with jBricks - A Toolkit for Addressing HCI Issues in Visual Language Environments
- URL: http://zvtm.sf.net

5.3. MapMosaic

KEYWORDS: Geo-visualization - Data visualization

SCIENTIFIC DESCRIPTION: GIS software applications and other mapping tools enable users to correlate data from multiple layers and gain insight from the resulting visualizations. However, most of these applications only feature basic, monolithic layer compositing techniques. These techniques do not always support users effectively in their tasks, as we observed during interviews with GIS experts. MapMosaic is a novel approach based on dynamic visual compositing that enables users to interactively create and manipulate local composites of multiple vector and raster map layers, taking into account the semantics and attribute values of objects and fields in the composit-ing process. We evaluated MapMosaic's interaction model against that of QGIS (a widely-used desktop GIS) and MAPublisher (a professional cartography tool) using the "Cognitive Dimensions" framework and through an analytical comparison, showing that MapMosaic's model is more flexible and can support users more effectively in their tasks.
FUNCTIONAL DESCRIPTION: MapMosaic is a novel approach to combine geographical layers based on dynamic visual compositing that enables users to interactively create and manipulate local composites of multiple vector and raster map layers. It takes into account the semantics and attribute values of objects and fields in the compositing process. MapMosaic aims at better supporting GIS users in their tasks such as correlating data from multiple layers and gaining insight from the resulting visualizations.

RELEASE FUNCTIONAL DESCRIPTION: First public release.

- Participants: Maria Jesus Lobo Gunther, Caroline Appert and Emmanuel Pietriga
- Contact: Emmanuel Pietriga
- Publications: MapMosaic: Dynamic Layer Compositing for Interactive Geovisualization - An Evaluation of Interactive Map Comparison Techniques
- URL: http://ilda.saclay.inria.fr/mapmuxing/mapmosaic/index.html

5.4. Baia

Before-and-after satellite image animation
KEYWORDS: Geo-visualization - 2D animation
SCIENTIFIC DESCRIPTION: Before-and-after image pairs show how entities in a given region have evolved over a specific period of time. Satellite images are a major source of such data, that capture how natural phenomena or human activity impact a geographical area. These images are used both for data analysis and to illustrate the resulting findings to diverse audiences. The simple techniques used to display them, including juxtaposing, swapping and monolithic blending, often fail to convey the underlying phenomenon in a meaningful manner. Baia is a framework to create advanced animated transitions, called animation plans, between before-and-after images. Baia relies on a pixel-based transition model that gives authors much expressive power, while keeping animations for common types of changes easy to create thanks to predefined animation primitives.
FUNCTIONAL DESCRIPTION: Baia is a framework to create advanced animated transitions, called animation plans, between before-and-after satellite images.

Before-and-after image pairs show how entities in a given region have evolved over a specific period of time. Satellite images are a major source of such data, that capture how natural phenomena or human activity impact a geographical area. These images are used both for data analysis and to illustrate the resulting findings to diverse audiences. The simple techniques used to display them, including juxtaposing, swapping and monolithic blending, often fail to convey the underlying phenomenon in a meaningful manner.

Baia relies on a pixel-based transition model that gives authors much expressive power. The animation editor enables authors to easily represent common types of changes thanks to predefined animation primitives and to sequence different changes across time.

RELEASE FUNCTIONAL DESCRIPTION: First public release

- Participants: Maria Jesus Lobo Gunther, Caroline Appert and Emmanuel Pietriga
- Contact: Emmanuel Pietriga
- Publication: Animation Plans for Before-and-After Satellite Images
- URL: http://ilda.saclay.inria.fr/mapmuxing/baia/index.html

5.5. LODAtlas

KEYWORDS: LOD - Linked open data - Semantic Web

SCIENTIFIC DESCRIPTION: The Web of Data is growing fast, as exemplified by the evolution of the Linked Open Data (LOD) cloud over the last ten years. One of the consequences of this growth is that it is becoming increasingly difficult for application developers and end-users to find the datasets that would be relevant to them. Semantic Web search engines, open data catalogs , datasets and frameworks such as LODStats and LOD Laundromat, are all useful but only give partial, even if complementary, views on what datasets are available on the Web. LODAtlas is a portal that enables users to find datasets of interest. Users can make different types of queries about both the datasets' metadata and contents, aggregated from multiple sources. They can then quickly evaluate the matching datasets' relevance, thanks to LODAtlas' summary visualizations of their general metadata, connections and contents.

FUNCTIONAL DESCRIPTION: The Web of Data is growing fast, as exemplified by the evolution of the Linked Open Data (LOD) cloud over the last ten years. One of the consequences of this growth is that it is becoming increasingly difficult for application developers and end-users to find the datasets that would be relevant to them. Semantic Web search engines, open data catalogs, datasets and frameworks such as LODStats and LOD Laundromat, are all useful but only give partial, even if complementary, views on what datasets are available on the Web. LODAtlas is a portal that enables users to find datasets of interest. Users can make different types of queries about both the datasets' metadata and contents, aggregated from multiple sources. They can then quickly evaluate the matching datasets' relevance, thanks to LODAtlas' summary visualizations of their general metadata, connections and contents.

- Participants: Caroline Appert, Marie Destandau, Ioana Manolescu, François Goasdoué, Sejla Cebiric, Hande Gozukan and Emmanuel Pietriga
- Contact: Emmanuel Pietriga
- Publication: Browsing Linked Data Catalogs with LODAtlas
- URL: http://lodatlas.lri.fr

5.6. TouchTokens

KEYWORDS: Tangible interface - HCI

SCIENTIFIC DESCRIPTION: TouchTokens make it possible to easily build interfaces that combine tangible and gestural input using passive tokens and a regular multi-touch surface. The tokens constrain users' grasp, and thus, the relative spatial configuration of fingers on the surface, theoretically making it possible to design algorithms that can recognize the resulting touch patterns. See associated scientific articles below.

FUNCTIONAL DESCRIPTION: TouchTokens allow interface designers to build low-cost tangible interfaces. The technique consists in recognizing multi-touch patterns that are associated with specific passive tokens. Those physical tokens can be made out of any material to get tracked on any touch-sensitive surface. Implementations of the recognizer (in both TUIO and Android) and vector descriptions of the tokens ready for 3D-printing or laser-cutting are available

- Participants: Caroline Appert, Rafael Morales Gonzalez, Emmanuel Pietriga and Gilles Bailly
- Contact: Caroline Appert
- Publications: TouchTokens: Guiding Touch Patterns with Passive Tokens - Passive yet Expressive TouchTokens - Custom-made Tangible Interfaces with TouchTokens
- URL: https://www.lri.fr/~appert/touchtokens/

5.7. Platforms

5.7.1. Platform: WILDER

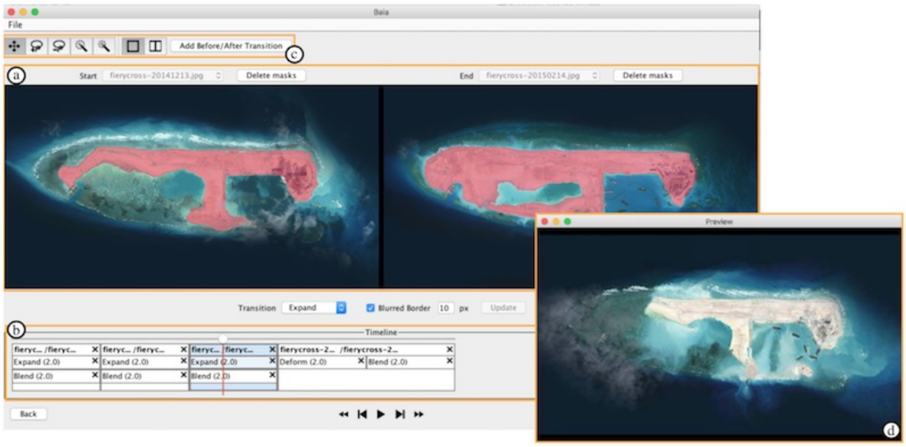


Figure 2. Baia is a framework to create advanced animated transitions between before-and-after satellite images. Before-and-after image pairs show how entities in a given region have evolved over a specific period of time. Baia relies on a pixel-based transition model to convey the underlying phenomenon that caused this evolution in a meaningful manner. The animation editor pictured here enables authors to easily represent common types of changes thanks to predefined animation primitives and to sequence different changes across time.

Figure 3. Multiple asteroid-generated tsunami simulations running simultaneously on the WILDER ultra-wall. The high display capacity of this interactive surface makes it possible to show, for each of the simulations: a planet-wide view showing the propagation of the tsunami on the globe, a close-up on the region of impact, showing a simulation of one or more scalar fields, parameters of the simulation.

Ultra-high-resolution wall-sized displays [39] feature a very high pixel density over a large physical surface. Such platforms have properties that make them well-suited to the visualization of very large datasets. They can represent the data with a high level of detail while at the same time retaining context: users can transition from an overview of the data to a detailed view simply by physically moving in front of the wall display. Wall displays also offer good support for collaborative work, enabling multiple users to simultaneously visualize and interact with the displayed data. To make them interactive, wall-sized displays are increasingly coupled with input devices such as touch frames, motion-tracking systems and wireless multitouch devices, in order to enable multi-device and multi-user interaction with the displayed data. Application areas for such visualization platforms range from the monitoring of complex infrastructures and crisis management situations to tools for the exploratory visualization of scientific data.

WILDER is the latest ultra-high-resolution wall-sized display set up at Inria Saclay, and is one of the nodes of the Digiscope EquipEx. We use this platform for multiple projects, both fundamental HCI research, and research and development activities for specific application areas such as geographical informations systems (Figure 3) and astronomy.

WILDER was used in the projects that led to the following publications this year: [23], [31].

6. New Results

6.1. Gestures and Tangibles

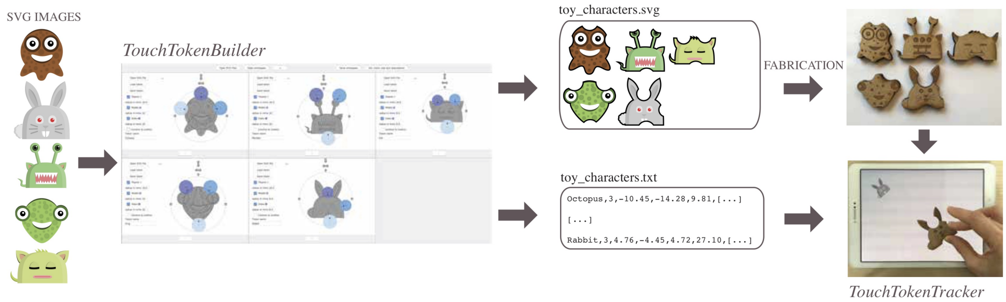


Figure 4. TouchToken-Builder (left) assists users in placing grasping notches on arbitrarily-shaped tokens, warning them about spatial configurations that could generate recognition conflicts or that might be uncomfortable to manipulate. It outputs both a vector and a numerical description of the tokens' geometry (middle). Those are used respectively to build the tokens (top-right), and to track them on any touchscreen using TouchToklen-Tracker (bottom-right).

6.1.1. Custom-made Tangible Interfaces with TouchTokens

One of our main results in this area is the design, development and evaluation of TouchTokens, a new way of prototyping and implementing low-cost tangible interfaces [6]. The approach requires only passive tokens and a regular multi-touch surface. The tokens constrain users' grasp, and thus, the relative spatial configuration of fingers on the surface, theoretically making it possible to design algorithms that can recognize the resulting touch patterns. Our latest project on TouchTokens [17] has been about tailoring tokens, going beyond the limited set of geometrical shapes studied in [6], as illustrated in Figure 4.

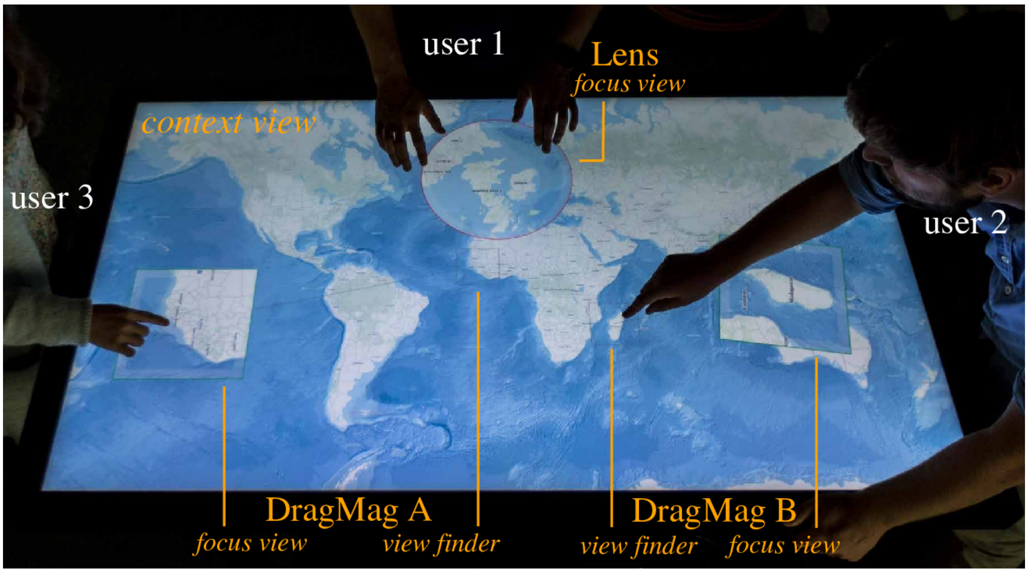


Figure 5. Our framework for the study of multi-scale navigation on tabletops enables users to both pan & zoom the context view and to create independent focus views, either DragMags or lenses.

6.1.2. Designing Coherent Gesture Sets for Multi-scale Navigation on Tabletops

We designed a framework for the study of multi-scale navigation (Figure 5) and conducted a controlled experiment of multi-scale navigation on tabletops [25]. We first conducted a guessability study in which we elicited user-defined gestures for triggering a coherent set of navigation actions, and then proposed two interface designs that combine the now-ubiquitous slide, pinch and turn gestures with either two-hand variations on these gestures, or with widgets. In a comparative study, we observed that users can easily learn both designs, and that the gesture-based, visually-minimalist design is a viable option, that saves display space for other controls.

6.1.3. Command Memorization, Gestures and other Triggering Methods

In collaboration with Telecom ParisTech, we studied the impact of semantic aids on command memorization when using either on-body interaction or directional gestures [21]. Previous studies had shown that spatial memory and semantic aids can help users learn and remember gestural commands. Using the body as a support to combine both dimensions had therefore been proposed, but no formal evaluations had been reported. We compared, with or without semantic aids, a new on-body interaction technique (BodyLoci) to mid-air Marking menus in a virtual reality context, considering three levels of semantic aids: no aid, story-making, and story-making with background images.

As part of the same collaboration, we also studied how memorizing positions or directions affects gesture learning for command selection. Many selection techniques either rely on directional gestures (e.g. Marking menus) or pointing gestures using a spatially-stable arrangement of items (e.g. FastTap). Both types of techniques are known to leverage memorization, but not necessarily for the same reasons. We investigated whether using directions or positions affects gesture learning [20].

6.2. Interacting with the Semantic Web of Linked Data

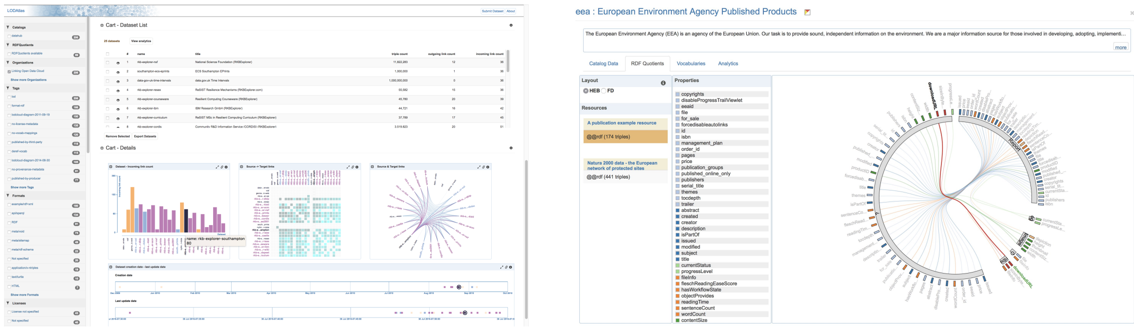


Figure 6. Browsing Linked Data Catalogs with LODAtlas [22]. Left: Visualization of the characteristics of, and links between, datasets selected by the user. Right: RDFQuotients-derived visual summary of a dataset. The summary shows how properties relate instances of the different classes.

The Web of Data is growing fast, as exemplified by the evolution of the Linked Open Data (LOD) cloud over the last ten years. One of the consequences of this growth is that it is becoming increasingly difficult for application developers and end-users to find the datasets that would be relevant to them. Semantic Web search engines, open data catalogs, datasets and frameworks such as LODStats and LOD Laundromat, are all useful but only give partial, even if complementary, views on what datasets are available on the Web. We started working on a platform called LODAtlas in 2016. LODAtlas [22] is a portal that enables users to find datasets of interest (see Figure 6) . Users can make different types of queries about both the datasets' metadata and contents, aggregated from multiple sources. They can then quickly evaluate the matching datasets' relevance, thanks to summary visualizations of their general metadata, connections and contents. The latter has been developed in collaboration with project-team CEDAR, based on their recent work on RDF Quotients.

Linked Data is structured as a directed labeled graph, or more precisely as a multitude of such graphs, that can be interlinked and distributed over the World Wide Web. Graph structures play an essential role at different scales in the Web of Data, and while it is now clear that basic approaches based on node-link diagram representations are only useful for small datasets, such visualizations remain meaningful for the representation of subsets of these multi-variate data. As part of a larger effort that started in the summer of 2016 to investigate novel interactive visual exploration techniques for multi-variate graphs, we introduced a design space and Web-based framework for generating what we call *animated edge textures*. Network edge data attributes are usually encoded using color, opacity, stroke thickness and stroke pattern, or some combination thereof. But in addition to these static variables, it is also possible to animate dynamic particles flowing along the edges. These can be seen as animated edge textures, that offer additional visual encodings that have potential not only in terms of visual mapping capacity but also playfulness and aesthetics. While such particle-based visual encodings have been featured in several commercial and design-oriented visualizations, this has to our knowledge almost always been done in a relatively ad hoc manner. Beyond the design space and Web framework, we also conducted an initial evaluation of particle properties – particle speed, pattern and frequency – in terms of visual perception. This work [24] was performed in collaboration with Nathalie Henry-Riche from Microsoft Research and Benjamin Bach from Edimburgh University.

6.3. Visualization

A significant part of our activity in this axis has been dedicated to geovisualization for various surfaces, including desktop workstations, tabletops and wall displays, in the context of ANR project MapMuxing. We investigated the representation of time in geovisualizations, more particularly how to convey changes in satellite images. Before-and-after images show how entities in a given region have evolved over a specific

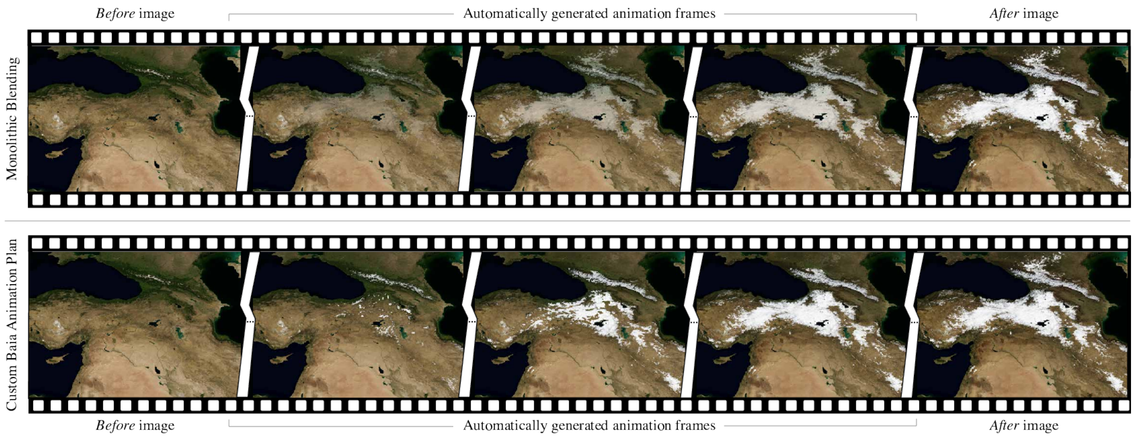


Figure 7. Animated transitions [4] based on one single before-and-after image pair showing seasonal snow cover over northern Middle East. The top row shows keyframes generated using basic monolithic blending. Snow fades in gradually but uniformly, regardless of altitude. The bottom row shows keyframes generated using a Baia animation plan derived from a Digital Elevation Model. Snow fades in gradually, but this time spreading from high-altitude to low-altitude areas.

period of time. These images are used both for data analysis and to illustrate the resulting findings to diverse audiences. We introduced Baia [4], a framework to create advanced animated transitions, called animation plans, between before-and-after images. Baia relies on a pixel-based transition model that gives authors much expressive power, while keeping animations for common types of changes easy to create thanks to predefined animation primitives (Figures 7 and 2).

Still in the area of geovisualization, in the context of ADT project Seawall, conducted in collaboration with project-team Lemon at Inria SAM / Montpellier and with Inria Chile, we have participated to the 2018 SciVis contest, which this year was about the visualization of data related to tsunamis generated by the impact of asteroids in deep water [31]. We used the WILDER ultra-high-resolution wall display to make it easier for analysts to visually compare and contrast different simulations from a deep water asteroid impact ensemble dataset. See Section 5.7.1 and Figure 3.

In the area of scientific data analysis, we have been collaborating with neuroscientists that explore large quantities of EEG data at different temporal scales. As a first step, we explored if automated algorithmic processes, that aid in the search for similar patterns in large datasets, actually match human intuition. We studied if we perceive as similar the results of these automatic measures, using three time-series visualizations: line charts, horizon graphs and colorfields. Our findings [15], [30] indicate that the notion of similarity is visualization-dependent, and that the best visual encoding varies depending on the automatic similarity measure considered.

Anstasia Bezerianos co-advised the PhD work of Evanthia Dimara in project-team Aviz together with P. Dragicevic. Last year, they had already confirmed that the cognitive bias known as the *attraction effect* does exist in visualizations [49]. This was followed-up this year by an exploration of different ways to mitigate this bias [12] (in collaboration with Northwestern University and Sorbonne Université). It was observed that the approach that consists of deleting all unwanted alternatives interactively removed the bias, a result that previous research has shown to be extremely hard to achieve. They also explored how different interactive visualizations of multidimensional datasets can affect decision making [13], and created a task-based taxonomy of cognitive biases for information visualization [14].

Our collaboration with INRA researchers has focused on mixed-initiative systems that combine human learning, machine learning and evolution. Results in this area for this year include an interactive evolutionary algorithm to learn from user interactions and steer the exploration of multidimensional datasets towards two-dimensional projections that are interesting to the analyst, and guidelines on how to evaluate such mixed initiative systems [29].

6.4. Collaboration, Multi-display environments, Large and Small Displays

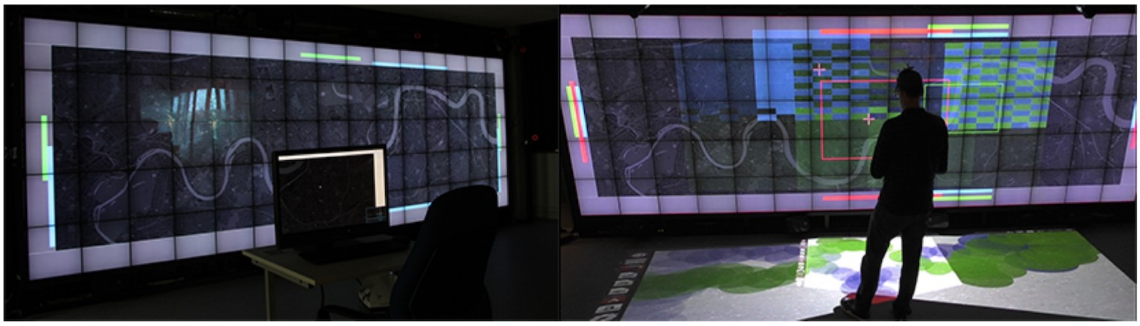


Figure 8. (left) Multi-Display Environment composed of a wall display and two workstations (one visible in the photo). (right) Three workspace awareness techniques: Awareness Bars at the edges of the wall, Focus Map on the wall display, and Step Map projected on the ground.

We studied awareness techniques to aid transitions between personal and shared workspaces in multi-display environments, that include large shared displays and desktops (Figure 8). In such contexts, including crisis management and control rooms, users can engage in both close collaboration and parallel or personal work. Transitioning between different displays can be challenging. To provide workspace awareness and to facilitate these transitions, we designed and implemented three interactions techniques that display users' activities. We explored how and where to display this activity: briefly on the shared display, or more persistently on a peripheral floor display. In a user study motivated by the context of a crisis room where multiple operators with different roles need to cooperate, we tested the usability of the techniques and provided insights on such transitions in systems running on MDEs [23]. We also contributed on a book chapter discussing how to best support collaboration in immersive environments that can range from MDE to mixed reality ones [28].

We collaborated with members from Inria project-team Aviz on the topic of small-scale visualization. This year, new results include a study about the perception of visualizations on smartwatches, performed together with Microsoft Research [11], [26]. The study was designed to assess how quickly people can perform a simple data comparison task for small-scale visualizations on a smartwatch. The goal was to extend our understanding of design constraints for smartwatch visualizations. We tested three chart types common on smartwatches: bar charts, donut charts, and radial bar charts with three different data sizes: 7, 12, and 24 data values. Results show that bar and donut charts should be preferred on smartwatch displays when quick data comparisons are necessary.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

- Tecknowmetrix (TKM): ANRT/CIFRE PhD (Hugo Romat), 3 years, started June 2016.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

MapMuxing - Multi-dimensional Map Multiplexing. (2014-2018) Funded by the French National Research Agency (ANR). In collaboration with IGN (Institut National de l'Information Géographique et Forestière): **208Keuros**/499Keuros. Participants: Emmanuel Pietriga (PI), Caroline Appert, Olivier Chapuis. http://mapmuxing.ign.fr

The project explores novel ways of combining different maps and data layers into a single cartographic representation, and investigates novel interaction techniques for navigating in it. The project aims at going beyond the traditional pan & zoom and overview+detail interface schemes, and at designing and evaluating novel cartographic visualizations that rely on high-quality generalization, *i.e.*, the simplification of geographic data to make it legible at a given map scale, and symbol specification.

8.1.2. Inria - Ministère de la Culture

Visual Exploration of Linked Data on BnF's data portal (2017-2018) Funded by the French Ministère de la Culture and Inria. **65Keuros**. Participants: Emmanuel Pietriga (PI), Caroline Appert, Hande Gözükan, Marie Destandau, Léo Colombaro.

The project explores novel ways of visually navigating the data exposed by the Bibliothèque Nationale de France as linked data on http://data.bnf.fr.

8.1.3. Inria Project Lab (IPL)

ILDA participates to Inria Project Lab iCODA : Data Journalism : knowledge-mediated Content and Data Interactive Analytics, that started in 2017. A key issue in data science is the design of algorithms that enable analysts to infer information and knowledge by exploring heterogeneous information sources, structured data, or unstructured content. With journalism data as a landmark use-case, iCODA aims to develop the scientific and technological foundation for collaborative, heterogeneous data analysis, guided by formalized, user-centric knowledge. The project relies on realistic scenarios in data-journalism to assess the contribution of the project to this area. iCODA is at the crossroads of several research areas (content analysis, data management, knowledge representation, visualization) and is part of a club of partners of the world of the press. Equipes-projets Inria : Graphik, Ilda, Linkmedia, Cedar. Press partners: Le Monde, OuestFrance, AFP. Participants: Anastasia Bezerianos (PI), Emmanuel Pietriga, Tong Xue, Nicole Barbosa Sultanum.

8.1.4. CNRS - PEPS

VizGest. (2018) Funded by CNRS. In collaboration with LIMSI. **17Keuros**. Participants: C. Appert (PI).

Interacting with multi-display environments often involves using mid-air gestures that do not require any proximity between users and displays. However, mid-air gestures are not *visible* to users. VizGest aims at giving some visibility to mid-air gestures by means of annotations put in the physical environment thanks to augmented reality glasses.

8.2. European Initiatives

8.2.1. Collaborations with Major European Organizations

Deutsches Elektronen-Synchrotron (DESY): Scientific collaboration on the design and implementation of user interfaces for array operations monitoring and control for the Cherenkov Telescope Array (CTA) project, to be built in the Canary Islands (Spain) and in the Atacama desert (Chile), 2 years, contract started May 2018

8.3. International Initiatives

8.3.1. Inria International Labs

Inria Chile. From 2012 to 2015, Emmanuel Pietriga was the scientific leader of the Massive Data team at Inria Chile, working on projects in collaboration with the ALMA radio-telescope and the Millenium Institute of Astrophysics. He is now scientific advisor to Inria Chile's visualization projects, and is actively involved in the collaboration between Inria Chile and the LSST on the design and development of user interfaces for operations monitoring and control (see below), and the project between ESO and Inria Chile about the design and implementation of user interfaces for ALMA's Integrated Alarm System.

8.3.2. Inria International Partners

Association of Universities for Research in Astronomy (AURA): contract, jointly with Inria Chile, on the design and implementation of user interfaces for telescope operations monitoring and control for the Large Synoptic Survey Telescope (LSST) project, under construction in the Atacama desert (Chile), started 2017. Participants: Emmanuel Pietriga (ILDA), José Galaz (Inria Chile), Sebastian Pereira (Inria Chile), Grazia Prato (Inria Chile).

8.4. International Research Visitors

8.4.1. Visits of International Scientists

8.4.1.1. Internships

- Nicole Barbosa Sultanum, Univ. Toronto, Canada, Oct 2018-Jan 2019.
- José Galaz, María Grazia Prato,Sebastian Pereira, Inria Chile, Dec 2018.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

- EICS 2018, 10th ACM SIGCHI Symposium on Engineering Interactive Computing Systems: Emmanuel Pietriga (general co-chair)

9.1.1.2. Member of the Organizing Committees

- VIS 2019, the IEEE Visualization Conference (SciVis, InfoVis, VAST): Anastasia Bezerianos (Workshops co-chair)
- VIS 2018, the IEEE Visualization Conference (SciVis, InfoVis, VAST): Anastasia Bezerianos (Communities co-chair)
- EICS 2018, 10th ACM SIGCHI Symposium on Engineering Interactive Computing Systems: Caroline Appert (Doctoral Consortium co-chair)
- IHM 2018, 30ème Conférence Francophone sur l'Interaction Homme-Machine: Caroline Appert (Doctoral Consortium panel member)

9.1.2. Scientific Events Selection

9.1.2.1. Chair of Conference Program Committees

- CHI 2019, 37th ACM SIGCHI Conference on Human Factors in Computing Systems: Anastasia Bezerianos (SC - subcommittee chair)
- CHI 2018, 36th ACM SIGCHI Conference on Human Factors in Computing Systems: Emmanuel Pietriga (SC - subcommittee chair)

9.1.2.2. Member of the Conference Program Committees

- CHI 2018, 36th ACM SIGCHI Conference on Human Factors in Computing Systems: Anastasia Bezerianos
- TheWebConf (WWW) 2019, 28th Web Conference, research track Web Content Analysis, Semantics, and Knowledge: Emmanuel Pietriga
- ESWC 2019, 16th Extended Semantic Web Conference: Emmanuel Pietriga
- ISWC 2018, 17th International Semantic Web Conference: Emmanuel Pietriga
- EICS 2018, 10th ACM SIGCHI Symposium on Engineering Interactive Computing Systems: Emmanuel Pietriga
- EICS 2018 TechNote, 10th ACM SIGCHI Symposium on Engineering Interactive Computing Systems: Caroline Appert, Olivier Chapuis
- VOILA @ ISWC 2018, Visualizations and User Interfaces for Ontologies and Linked Data, workshop co-located with ISWC 2018: Emmanuel Pietriga
- TheWebConf (WWW) 2018, 27th Web Conference, research track Web Content Analysis, Semantics, and Knowledge: Emmanuel Pietriga
- IHM 2018, 30ème Conférence Francophone sur l'Interaction Homme-Machine: Caroline Appert

9.1.2.3. Reviewer

- ACM CHI 2019, Conference on Human Factors in Computing Systems: Caroline Appert, Olivier Chapuis, Emmanuel Pietriga, Hugo Romat
- ACM UIST 2018, Interface Software and Technologies Symposium: Olivier Chapuis, Caroline Appert
- ACM ISS 2018, International Conference on Interactive Surfaces and Spaces: Olivier Chapuis
- AVI 2018, Advanced Visual Interfaces: Olivier Chapuis
- IEEE VIS 2018, Visualization Conference (InfoVis): Anastasia Bezerianos
- IHM 2018, Conference of the Association Francophone d'Interaction Homme-Machine: Olivier Chapuis, Bruno Fruchard, Marie Destandau

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- ACM ToCHI, Transactions on Computer-Human Interaction: Caroline Appert (associate editor and member of EiC search committee)

9.1.3.2. Reviewer - Reviewing Activities

- IEEE TVCG, Transactions on Visualization and Computer Graphics: Emmanuel Pietriga, Anastasia Bezerianos
- Information Visualization Journal: Emmanuel Pietriga
- Journal of Multimodal Interfaces: Caroline Appert
- Interacting with Computers: Caroline Appert

9.1.4. Invited Talks

- Olivier Chapuis: Expressive and multi-user interaction, desktops, wall displays and beyond, Dresden Talks on Interaction & Visualization, Technische Universität Dresden (Germany), May 2018.
- Caroline Appert: Low-cost Tangible Interaction with TouchTokens, Conference FAB' 14, Paris Saclay, July 2018.
- Anastasia Bezerianos: How interactive visualizations can aid us make better or at least more consistent choices, University of Calgary (Canada), Nov 2018.

- Anastasia Bezerianos: Visualization Perception at large and small scales, University of St. Andrews (UK), Dec 2018.

9.1.5. Leadership within the Scientific Community

- ANR, CES chair - Interaction, Robotique: Caroline Appert

9.1.6. Scientific Expertise

- H2020, ERC Advanced Grants (reviewer): Caroline Appert, Emmanuel Pietriga
- ANR, CES chair - Interaction, Robotique: Caroline Appert
- NSERC, Canadian Discovery Grands (reviewer): Anastasia Bezerianos

9.1.7. Research Administration

- Scientific coordinator of Inria's evaluation seminar for period 2015-2018, theme Interaction and visualization: Emmanuel Pietriga
- Deputy Director of the Laboratoire de Recherche en Informatique (LRI): Olivier Chapuis
- President of Inria Saclay - Île de France's Commission for Technological Development (CDT): Emmanuel Pietriga
- Deputy head of Pôle "Données, Connaissances, Apprentissage et Interaction" at École Doctorale Paris Saclay (ED STIC): Caroline Appert

9.1.8. Learned societies

- Association Francophone d'Interaction Homme-Machine (AFIHM), in charge of the relation with the SIF: Olivier Chapuis.
- SigCHI Paris Local Chapter, chair: Anastasia Bezerianos.
- SigCHI Paris Local Chapter, vice chair: Caroline Appert.

9.1.9. Hiring committees

- Univ. Paris-Sud hiring committee, Commission Consultative des Spécialistes de l'Université 27ème section (computer science), members: Caroline Appert, Anastasia Bezerianos.
- Assistant Professor position, Computer Science, Université Paris-Sud, 2018: Caroline Appert, Emmanuel Pietriga.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master Co-head: Anastasia Bezerianos, M2 Interaction and HCID, Univ. Paris-Saclay (until Aug 2018).

Ingénieur: Emmanuel Pietriga, Data Visualization (INF552), 36h, 3A/M1, École Polytechnique.

Ingénieur: Caroline Appert, Data Visualization (INF552), 18h, 3A/M1, École Polytechnique.

Master: Emmanuel Pietriga, Data Visualization, 24h, M2 Informatique Décisionnelle, Univ. Paris-Dauphine.

Master: Caroline Appert, Evaluation of Interactive Systems - Introduction, 21h, M2 Interaction and HCID, Univ. Paris-Saclay.

Master: Caroline Appert, Evaluation of Interactive Systems - Advanced, 21h, M2 Interaction and HCID, Univ. Paris-Saclay.

Master: Anastasia Bezerianos, Introduction to Programming of Interactive Systems, 21h, M2 Interaction and HCID, Univ. Paris-Saclay.

Master: Anastasia Bezerianos, Career Seminar, 21h, M2 Interaction and HCID, Univ. Paris-Saclay.

Master: Anastasia Bezerianos, Mixed Reality and Tangible Interaction, 11h, M2 Interaction and HCID, Univ. Paris-Saclay.

Master: Anastasia Bezerianos, Information Visualization, 10h, M2 Interaction and HCID, Univ. Paris-Saclay.

Master: Anastasia Bezerianos, HCI Project, 21h, M2 Interaction and HCID, Univ. Paris-Saclay.

Master: Anastasia Bezerianos, Design Project, 21h, M1 HCID, Univ. Paris-Saclay.

Master: Anastasia Bezerianos, Introduction aux Systèmes Interactifs, 21h, M1 Informatique, Univ. Paris-Saclay.

Master: Raphaël James, Programming of Interactive Systems, 12h, Univ. Paris-Saclay

Master: Raphaël James, Fundamentals of HCI, 21h, Univ. Paris-Saclay

Master: Tong Xue, Design of Interactive Systems, 21h, Univ. Paris-Saclay

License: Raphaël James, Algorithm/C/C++, 29h, Polytech

License: Tong Xue, Introduction à l'informatique graphique, 33h, Univ. Paris-Sud

Licence: Bruno Fruchard, Visualisation, 10h, Télécom Paristech.

Licence: Bruno Fruchard, Interaction Homme-Machine, 8h, Télécom Paristech.

Licence: Bruno Fruchard, Paradigmes de programmation, 8h, Télécom Paristech.

Licence: Bruno Fruchard, Projet d'apprentissage collaboratif thématique (student supervision), 6h, Télécom Paristech.

IUT: Marie Destandau, Programmation et administration des bases de données, 32h, Univ. Paris-Sud

IUT: Marie Destandau, Suivi de projet, 27h, Univ. Paris-Sud

IUT: Marie Destandau, Bases de données avancées, 18h, Univ. Paris-Sud

9.2.2. Supervision

PhD in progress : Eugénie Brasier, Interaction techniques for remote manipulation in multi-display environments, since October 2018, Advisors: Caroline Appert

PhD in progress : Raphaël James, Environnements de réalité physique et augmentée utilisés dans l'analyse visuelle collaborative, since October 2018, Advisors: Anastasia Bezerianos, Olivier Chapuis, Tim Dwyer

PhD in progress : Tong Xue, Interactive Visualization for Data Journalism, since October 2018, Advisors: Anastasia Bezerianos, Emmanuel Pietriga

PhD in progress : Marie Destandau, Interactive Visual Exploration of Webs of Data, since October 2017, Advisors: Caroline Appert, Emmanuel Pietriga

PhD in progress : Anna Gogolou, Iterative and expressive querying for big data series, October 2016, Advisors: Anastasia Bezerianos, Themis Palpanas

PhD in progress : Hugo Romat, Visual exploration and interactive manipulation techniques for collections of heterogeneous data and documents, since June 2016, Advisors: Caroline Appert, Emmanuel Pietriga

PhD in progress : Bruno Fruchard, Techniques d'interaction exploitant la mémoire spatiale pour faciliter l'accès rapide aux commandes et aux données, since October 2015 – December 2018, Advisors: Eric Lecolinet, Olivier Chapuis

9.2.3. Juries

PhD: Kashyap Todi, Hasselt University, Belgium: Emmanuel Pietriga (rapporteur)

PhD: Anonymous, University of Swinburne, Australia: Emmanuel Pietriga (rapporteur)

PhD: Ulrich von Zadow PhD, Technische Universität Dresden, Germany: Olivier Chapuis (reviewer)

PhD: Juliette Rambourg, Université de Toulouse, France: Caroline Appert (rapporteur)

PhD: Julien Gori, Université Paris Saclay, France: Caroline Appert (président)

PhD: Marion Dumont, Université Paris-Est, France: Caroline Appert (examinateur)

PhD: Fatemeh Rajabiyazdi, University of Calgary, Canada: Anastasia Bezerianos (rapporteur)

PhD: Amira Chalbi, Université Lille 1, France: Anastasia Bezerianos (examinateur)

HDR: Marcos Serrano, Université de Toulouse, France: Caroline Appert (président)

10. Bibliography

Major publications by the team in recent years

[1] C. Appert, O. Chapuis, E. Pietriga, M.-J. Lobo.*Reciprocal Drag-and-Drop*, in "ACM Transactions on Computer-Human Interaction", September 2015, vol. 22, n^{o} 6, p. 29:1–29:36 [*DOI :* 10.1145/2785670], https://hal.archives-ouvertes.fr/hal-01185805

[2] O. Chapuis, A. Bezerianos, S. Frantzeskakis.*Smarties: An Input System for Wall Display Development*, in "CHI '14", Toronto, Canada, ACM, April 2014, p. 2763-2772 [*DOI :* 10.1145/2556288.2556956], https://hal.archives-ouvertes.fr/hal-00979034

[3] C. Liu, O. Chapuis, M. Beaudouin-Lafon, E. Lecolinet.*CoReach: Cooperative Gestures for Data Manipulation on Wall-sized Displays*, in "Proceedings of the 35th international conference on Human factors in computing systems", Denver, United States, CHI '17, ACM, May 2017 [*DOI :* 10.1145/3025453.3025594], https://hal.archives-ouvertes.fr/hal-01437091

[4] M.-J. Lobo, C. Appert, E. Pietriga.*Animation Plans for Before-and-After Satellite Images*, in "IEEE Transactions on Visualization and Computer Graphics", January 2018, vol. 24 [*DOI :* 10.1109/TVCG.2018.2796557], https://hal.inria.fr/hal-01773882

[5] M.-J. Lobo, E. Pietriga, C. Appert.*An Evaluation of Interactive Map Comparison Techniques*, in "CHI '15 Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems", Seoul, South Korea, ACM, April 2015, p. 3573-3582 [*DOI :* 10.1145/2702123.2702130], https://hal.inria.fr/hal-01144163

[6] R. Morales González, C. Appert, G. Bailly, E. Pietriga.*TouchTokens: Guiding Touch Patterns with Passive Tokens*, in "2016 CHI Conference on Human Factors in Computing Systems", San Jose, CA, United States, May 2016 [*DOI :* 10.1145/2858036.2858041], https://hal.archives-ouvertes.fr/hal-01315130

[7] H. Olafsdottir, C. Appert.*Multi-Touch Gestures for Discrete and Continuous Control*, in "International Working Conference on Advanced Visual Interfaces (AVI)", Como, Italy, May 2014, 8 [*DOI :* 10.1145/2598153.2598169], https://hal.archives-ouvertes.fr/hal-00998971

[8] E. Pietriga, F. Del Campo, A. Ibsen, R. Primet, C. Appert, O. Chapuis, M. Hempel, R. Muñoz, S. Eyheramendy, A. Jordan, H. Dole.*Exploratory Visualization of Astronomical Data on Ultra-high-resolution Wall Displays*, in "Proceedings SPIE", July 2016, vol. 9913, 15 [*DOI :* 10.1117/12.2231191], https://hal.inria.fr/hal-01350722

[9] A. PROUZEAU, A. BEZERIANOS, O. CHAPUIS.*Evaluating Multi-User Selection for Exploring Graph Topology on Wall-Displays*, in "IEEE Transactions on Visualization and Computer Graphics", August 2017, vol. 23, n^{o} 8, p. 1936–1951 [*DOI :* 10.1109/TVCG.2016.2592906], https://hal.archives-ouvertes.fr/hal-01348578

[10] T. TSANDILAS, A. BEZERIANOS, T. JACOB.*SketchSliders: Sketching Widgets for Visual Exploration on Wall Displays*, in "Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems", Seoul, South Korea, ACM, April 2015, p. 3255-3264 [*DOI :* 10.1145/2702123.2702129], https://hal.archives-ouvertes.fr/hal-01144312

Publications of the year

Articles in International Peer-Reviewed Journal

[11] T. BLASCHECK, L. BESANÇON, A. BEZERIANOS, B. LEE, P. ISENBERG.*Glanceable Visualization: Studies of Data Comparison Performance on Smartwatches*, in "IEEE Transactions on Visualization and Computer Graphics", January 2019, vol. 25, n^{o} 1, https://hal.inria.fr/hal-01851306

[12] E. DIMARA, G. BAILLY, A. BEZERIANOS, S. FRANCONERI.*Mitigating the Attraction Effect with Visualizations*, in "IEEE Transactions on Visualization and Computer Graphics", October 2018, 11 [*DOI :* 10.1109/TVCG.2018.2865233], https://hal.inria.fr/hal-01845004

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

[14] E. DIMARA, S. FRANCONERI, C. PLAISANT, A. BEZERIANOS, P. DRAGICEVIC.*A Task-based Taxonomy of Cognitive Biases for Information Visualization*, in "IEEE Transactions on Visualization and Computer Graphics", 2019, https://hal.sorbonne-universite.fr/hal-01868738

[15] A. GOGOLOU, T. TSANDILAS, T. PALPANAS, A. BEZERIANOS.*Comparing Similarity Perception in Time Series Visualizations*, in "IEEE Transactions on Visualization and Computer Graphics", October 2018, 11, https://hal.inria.fr/hal-01845008

[16] M.-J. LOBO, C. APPERT, E. PIETRIGA.*Animation Plans for Before-and-After Satellite Images*, in "IEEE Transactions on Visualization and Computer Graphics", January 2018, vol. 24 [*DOI :* 10.1109/TVCG.2018.2796557], https://hal.inria.fr/hal-01773882

International Conferences with Proceedings

[17] C. APPERT, E. PIETRIGA, E. BARTENLIAN, R. MORALES GONZÁLEZ.*Custom-made Tangible Interfaces with TouchTokens*, in "Proceedings of the International Working Conference on Advanced Visual Interfaces", Grosseto, Italy, AVI '18, ACM, May 2018 [*DOI :* 10.1145/3206505.3206509], https://hal.archives-ouvertes.fr/hal-01777599

[18] T. BLASCHECK, A. BEZERIANOS, L. BESANÇON, B. LEE, P. ISENBERG.*Preparing for Perceptual Studies: Position and Orientation of Wrist-worn Smartwatches for Reading Tasks*, in "Proceedings of the Workshop on Data Visualization on Mobile Devices held at ACM CHI", Montréal, Canada, April 2018, https://hal.inria.fr/hal-01744246

[19] B. FRUCHARD, E. LECOLINET, O. CHAPUIS.*Command Memorisation: Spatial Positions versus Directional Gestures*, in "30eme conférence francophone sur l'interaction homme-machine", Brest, France, Articles Scientifiques, AFIHM, October 2018, p. 92-99, https://hal.archives-ouvertes.fr/hal-01899048

[20] B. FRUCHARD, E. LECOLINET, O. CHAPUIS.*How Memorizing Positions or Directions Affects Gesture Learning?*, in "Proceedings of the 2018 International Conference on Interactive Surfaces and Spaces", Tokyo, Japan, ISS '18, ACM, November 2018, p. 107–114 [*DOI :* 10.1145/3279778.3279787], https://hal.archives-ouvertes.fr/hal-01891436

[21] B. FRUCHARD, E. LECOLINET, O. CHAPUIS.*Impact of Semantic Aids on Command Memorization for On-Body Interaction and Directional Gestures*, in "AVI 2018 - International Conference on Advanced Visual Interfaces", Castiglione della Pescaia, Grosseto, Italy, ACM, May 2018 [*DOI :* 10.1145/3206505.3206524], https://hal.archives-ouvertes.fr/hal-01764757

[22] E. PIETRIGA, H. GÖZÜKAN, C. APPERT, M. DESTANDAU, Š. ČEBIRIĆ, F. GOASDOUÉ, I. MANOLESCU.*Browsing Linked Data Catalogs with LODAtlas*, in "ISWC 2018 - 17th International Semantic Web Conference", Monterey, United States, Springer, October 2018, p. 137-153, https://hal.inria.fr/hal-01827766

[23] A. PROUZEAU, A. BEZERIANOS, O. CHAPUIS.*Awareness Techniques to Aid Transitions between Personal and Shared Workspaces in Multi-Display Environments*, in "Proceedings of the 2018 International Conference on Interactive Surfaces and Spaces", Tokyo, Japan, ISS '18, ACM, November 2018, p. 291–304 [*DOI :* 10.1145/3279778.3279780], https://hal.archives-ouvertes.fr/hal-01890787

[24] H. ROMAT, C. APPERT, B. BACH, N. HENRY-RICHE, E. PIETRIGA.*Animated Edge Textures in Node-Link Diagrams: a Design Space and Initial Evaluation*, in "Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems", Montréal, Canada, CHI '18, ACM, April 2018, p. 187:1–187:13 [*DOI :* 10.1145/3173574.3173761], https://hal.inria.fr/hal-01726358

[25] V. RUSŇÁK, C. APPERT, O. CHAPUIS, E. PIETRIGA.*Designing Coherent Gesture Sets for Multi-scale Navigation on Tabletops*, in "Proceedings of the 36th international conference on Human factors in computing systems", Montreal, Canada, CHI '18, ACM, April 2018, p. 142:1-142:12 [*DOI :* 10.1145/3173574.3173716], https://hal.inria.fr/hal-01722189

Conferences without Proceedings

[26] T. BLASCHECK, L. BESANÇON, A. BEZERIANOS, B. LEE, P. ISENBERG.*Visualization perception on smartwatches*, in "Journée Visu 2018", Palaiseau, France, May 2018, https://hal.inria.fr/hal-01844726

[27] B. FRUCHARD, E. LECOLINET, O. CHAPUIS.*Command Memorization: Spatial Positions versus Directional Gestures*, in "Proceedings of the 30e Confe´rence Francophone sur l'Interaction Homme-Machine", Chauvigny, France, October 2018 [*DOI :* 10.1145/3286689.3286700], https://hal.archives-ouvertes.fr/hal-01898325

Scientific Books (or Scientific Book chapters)

[28] M. BILLINGHURST, M. CORDEIL, A. BEZERIANOS, T. MARGOLIS.*Collaborative Immersive Analytics*, in "Immersive Analytics", Marriott, K., Schreiber, F., Dwyer, T., Klein, K., Riche, N.H., Itoh, T., Stuerzlinger, W., Thomas, B.H., October 2018 [*DOI :* 10.1007/978-3-030-01388-2_8], https://hal.inria.fr/hal-01938529

[29] N. BOUKHELIFA, A. BEZERIANOS, E. LUTTON.*Evaluation of Interactive Machine Learning Systems*, in "Human and Machine LearningVisible, Explainable, Trustworthy and Transparent", J. ZHOU, F. CHEN (editors), Springer, July 2018, p. 341-360, https://hal.inria.fr/hal-01845018

Research Reports

[30] A. GOGOLOU, T. TSANDILAS, T. PALPANAS, A. BEZERIANOS.*Comparing Time Series Similarity Perception under Different Color Interpolations*, Inria, June 2018, n^{o} RR-9189, https://hal.inria.fr/hal-01844994

Other Publications

[31] A. KAMAKSHIDASAN, J. GALAZ, R. CIENFUEGOS, A. ROUSSEAU, E. PIETRIGA.*Comparative Visualization of Deep Water Asteroid Impacts on Ultra-high-resolution Wall Displays with Seawall*, October 2018, p. 1-2, IEEE VIS 2018 - IEEE Conference on Visualization, Poster, https://hal.inria.fr/hal-01888112

[32] A. KAMAKSHIDASAN.*Spatial Comparison of Cricketers*, October 2018, p. 1-2, IEEE VIS 2018 - IEEE Conference on Visualization, Poster, https://hal.inria.fr/hal-01852138

References in notes

[33] M. R. ENDSLEY, D. G. JONES (editors). *Designing for Situation Awareness: an Approach to User-Centered Design*, CRC Press, Taylor & Francis, 2012

[34] S. ABITEBOUL, P. BUNEMAN, D. SUCIU.*Data on the Web: From Relations to Semistructured Data and XML*, Morgan Kaufmann, 1999

[35] S. ABITEBOUL, P. SENELLART, V. VIANU.*The ERC Webdam on Foundations of Web Data Management*, in "Proc. of the 21st International Conference Companion on World Wide Web", WWW '12 Companion, ACM, 2012, p. 211–214, http://doi.acm.org/10.1145/2187980.2188011

[36] J. ALVINA, C. APPERT, O. CHAPUIS, E. PIETRIGA.*RouteLens: Easy Route Following for Map Applications*, in "Proceedings of the International Working Conference on Advanced Visual Interfaces", AVI '14, ACM, 2014, p. 125–128 [*DOI :* 10.1145/2598153.2598200]

[37] C. APPERT, O. CHAPUIS, E. PIETRIGA.*High-precision magnification lenses*, in "Proceedings of the 28th international conference on Human factors in computing systems", CHI '10, ACM, 2010, p. 273–282 [*DOI :* 10.1145/1753326.1753366]

[38] C. APPERT, J.-D. FEKETE.*OrthoZoom Scroller: 1D Multi-scale Navigation*, in "Proceedings of the SIGCHI Conference on Human Factors in Computing Systems", New York, NY, USA, CHI '06, ACM, 2006, p. 21–30, http://doi.acm.org/10.1145/1124772.1124776

[39] M. BEAUDOUIN-LAFON, S. HUOT, M. NANCEL, W. MACKAY, E. PIETRIGA, R. PRIMET, J. WAGNER, O. CHAPUIS, C. PILLIAS, J. R. EAGAN, T. GJERLUFSEN, C. KLOKMOSE.*Multisurface Interaction in the WILD Room*, in "IEEE Computer", 2012, vol. 45, n^{o} 4, p. 48-56 [*DOI :* 10.1109/MC.2012.110]

[40] E. BENSON, D. KARGER.*End-users Publishing Structured Information on the Web: An Observational Study of What, Why, and How*, in "Proc. of the SIGCHI Conference on Human Factors in Computing Systems", CHI '14, ACM, 2014, p. 1265–1274, http://doi.acm.org/10.1145/2556288.2557036

[41] J. BERNDTSSON, M. NORMARK.*The coordinative functions of flight strips: air traffic control work revisited*, in "Proceedings of the international ACM SIGGROUP conference on Supporting group work", GROUP '99, ACM, 1999, p. 101–110, http://doi.acm.org/10.1145/320297.320308

[42] T. BERNERS-LEE, Y. CHEN, L. CHILTON, D. CONNOLLY, R. DHANARAJ, J. HOLLENBACH, A. LERER, D. SHEETS.*Tabulator: Exploring and Analyzing linked data on the Semantic Web*, in "Proc. of the International Semantic Web User Interaction Workshop", 2006

[43] J. BERTIN.*Semiology of Graphics*, University of Wisconsin Press, 1983

[44] A. BEZERIANOS, R. BALAKRISHNAN.*View and Space Management on Large Displays*, in "IEEE Computer Graphics and Applications", 2005, vol. 25, n^o 4, p. 34-43

[45] A. BEZERIANOS.*Using alternative views for layout, comparison and context switching tasks in wall displays*, in "OZCHI", 2007, p. 303-310

[46] D. BONNET, C. APPERT, M. BEAUDOUIN-LAFON.*Extending the Vocabulary of Touch Events with ThumbRock*, in "Proceedings of Graphics Interface", GI '13, CIPS, 2013, p. 221-228, http://dl.acm.org/citation.cfm?id=2532129.2532166

[47] A. COLLINS, A. BEZERIANOS, G. MCEWAN, M. RITTENBRUCH, R. WASINGER, J. KAY.*Understanding File Access Mechanisms for Embedded Ubicomp Collaboration Interfaces*, in "Proceedings of the 11th International Conference on Ubiquitous Computing", New York, NY, USA, UbiComp '09, ACM, 2009, p. 135–144, http://doi.acm.org/10.1145/1620545.1620567

[48] A.-S. DADZIE, M. ROWE.*Approaches to Visualising Linked Data: A Survey*, in "Semantic Web", 2011, vol. 2, n^o 2, http://dx.doi.org/10.3233/SW-2011-0037

[49] E. DIMARA, A. BEZERIANOS, P. DRAGICEVIC.*The Attraction Effect in Information Visualization*, in "IEEE Transactions on Visualization and Computer Graphics", 2017, vol. 23, n^o 1 [*DOI :* 10.1109/TVCG.2016.2598594], https://hal.inria.fr/hal-01355750

[50] A. DIX, G. LEPOURAS, A. KATIFORI.*From the web of data to a world of action*, in "Web Semantics", 2010, vol. 8, p. 394–408, http://dx.doi.org/10.1016/j.websem.2010.04.007

[51] D. FENSEL, J. HENDLER, H. LIEBERMAN, W. WAHLSTER.*Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*, MIT Press, 2005

[52] C. FORLINES, A. ESENTHER, C. SHEN, D. WIGDOR, K. RYALL.*Multi-user, Multi-display Interaction with a Single-user, Single-display Geospatial Application*, in "Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology", UIST '06, ACM, 2006, p. 273–276, http://doi.acm.org/10.1145/1166253.1166296

[53] J. GRUDIN.*Computer-Supported Cooperative Work: History and Focus*, in "IEEE Computer", May 1994, vol. 27, n^o 5, p. 19–26, http://dx.doi.org/10.1109/2.291294

[54] T. HEATH, C. BIZER.*Linked Data: Evolving the Web into a Global Data Space*, Morgan & Claypool, 2011

[55] J. HEBELER, M. FISHER, R. BLACE, A. PEREZ-LOPEZ.*Semantic web programming*, Wiley & Sons, 2011

[56] M. HILDEBRAND, J. VAN OSSENBRUGGEN, L. HARDMAN./*facet: A Browser for Heterogeneous Semantic Web Repositories*, in "ISWC'06: Proc. of the International Semantic Web Conference", Springer, 2006, p. 272–285

[57] D. F. HUYNH, R. C. MILLER, D. KARGER.*Potluck: Data Mash-up Tool for Casual Users*, in "Web Semantics", 2008, vol. 6, n^{o} 4, p. 274–282, http://dx.doi.org/10.1016/j.websem.2008.09.005

[58] H. ISHII, B. ULLMER.*Tangible Bits: Towards Seamless Interfaces Between People, Bits and Atoms*, in "Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems", CHI '97, ACM, 1997, p. 234–241, http://doi.acm.org/10.1145/258549.258715

[59] J. A. JACKO.*Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications, 3rd Edition*, 3rd, CRC Press, 2012

[60] D. KARGER, S. OSTLER, R. LEE.*The Web Page As a WYSIWYG End-user Customizable Database-backed Information Management Application*, in "Proc. of the Symposium on User Interface Software and Technology", UIST '09, ACM, 2009, p. 257–260, http://doi.acm.org/10.1145/1622176.1622223

[61] A. KATIFORI, C. HALATSIS, G. LEPOURAS, C. VASSILAKIS, E. GIANNOPOULOU.*Ontology visualization methods—a survey*, in "ACM Computing Surveys", 2007, vol. 39, n^{o} 4, p. 10:1-10:42

[62] C. LIU, O. CHAPUIS, M. BEAUDOUIN-LAFON, E. LECOLINET.*Shared Interaction on a Wall-Sized Display in a Data Manipulation Task*, in "Proceedings of the 34th international conference on Human factors in computing systems", San Jose, United States, CHI '16, ACM, May 2016, p. 1-12 [*DOI :* 10.1145/2858036.2858039], https://hal.inria.fr/hal-01275535

[63] C. LIU, O. CHAPUIS, M. BEAUDOUIN-LAFON, E. LECOLINET, W. MACKAY.*Effects of Display Size and Navigation Type on a Classification Task*, in "Proceedings of the 32nd international conference on Human factors in computing systems", CHI '14, ACM, 2014, p. 4147-4156, http://doi.acm.org/10.1145/2556288.2557020

[64] W. MACKAY.*Is paper safer? The role of paper flight strips in air traffic control*, in "ACM Trans. Comput.-Hum. Interact.", December 1999, vol. 6, p. 311–340, http://doi.acm.org/10.1145/331490.331491

[65] R. MORALES GONZÁLEZ, C. APPERT, G. BAILLY, E. PIETRIGA.*TouchTokens: Guiding Touch Patterns with Passive Tokens*, in "Proceedings of the SIGCHI Conference on Human Factors in Computing Systems", New York, NY, USA, CHI '16, ACM, 2016, To appear

[66] T. MUNZNER.*Visualization Analysis and Design*, CRC Press, 2014

[67] M. NANCEL, O. CHAPUIS, E. PIETRIGA, X.-D. YANG, P. P. IRANI, M. BEAUDOUIN-LAFON.*High-precision pointing on large wall displays using small handheld devices*, in "Proceedings of the SIGCHI Conference on Human Factors in Computing Systems", CHI '13, ACM, 2013, p. 831–840, http://doi.acm.org/10.1145/2470654.2470773

[68] M. NANCEL, E. PIETRIGA, O. CHAPUIS, M. BEAUDOUIN-LAFON.*Mid-air Pointing on Ultra-Walls*, in "ACM Transactions on Computer-Human Interaction", August 2015, vol. 22, n^{o} 5, p. 21:1–21:62 [*DOI :* 10.1145/2766448], https://hal.inria.fr/hal-01184544

[69] A.-C. NGONGA NGOMO, L. BÜHMANN, C. UNGER, J. LEHMANN, D. GERBER.*Sorry, I Don'T Speak SPARQL: Translating SPARQL Queries into Natural Language*, in "Proc. of the International Conference on World Wide Web", WWW '13, ACM, 2013, p. 977–988, http://dl.acm.org/citation.cfm?id=2488388.2488473

[70] H. OLAFSDOTTIR, T. TSANDILAS, C. APPERT.*Prospective Motor Control on Tabletops: Planning Grasp for Multitouch Interaction*, in "Proceedings of the 32nd international conference on Human factors in computing systems", CHI '14, ACM, 2014, p. 2893-2902, http://doi.acm.org/10.1145/2556288.2557029

[71] E. PIETRIGA, C. APPERT, M. BEAUDOUIN-LAFON.*Pointing and beyond: an operationalization and preliminary evaluation of multi-scale searching*, in "CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems", ACM Press, 2007, p. 1215–1224 [*DOI :* 10.1145/1240624.1240808]

[72] E. PIETRIGA, C. APPERT.*Sigma lenses: focus-context transitions combining space, time and translucence*, in "CHI '08: Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems", ACM, 2008, p. 1343–1352 [*DOI :* 10.1145/1357054.1357264]

[73] E. PIETRIGA, O. BAU, C. APPERT.*Representation-Independent In-Place Magnification with Sigma Lenses*, in "IEEE Transactions on Visualization and Computer Graphics (TVCG)", 2010, vol. 16, n^{o} 03, p. 455-467 [*DOI :* 10.1109/TVCG.2009.98]

[74] E. PIETRIGA, C. BIZER, D. KARGER, R. LEE.*Fresnel - A Browser-Independent Presentation Vocabulary for RDF*, in "Proceedings of the 5th International Semantic Web Conference (ISWC 2006)", Springer, November 2006, p. 158-171

[75] E. PIETRIGA, S. HUOT, M. NANCEL, R. PRIMET.*Rapid development of user interfaces on cluster-driven wall displays with jBricks*, in "Proceedings of the 3rd ACM SIGCHI symposium on Engineering interactive computing systems", EICS '11, ACM, 2011, p. 185–190, http://doi.acm.org/10.1145/1996461.1996518

[76] E. PIETRIGA.*A Toolkit for Addressing HCI Issues in Visual Language Environments*, in "IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)", IEEE Computer Society, 2005, p. 145-152 [*DOI :* 10.1109/VLHCC.2005.11]

[77] C. PINDAT, E. PIETRIGA, O. CHAPUIS, C. PUECH.*JellyLens: Content-Aware Adaptive Lenses*, in "UIST - 25th Symposium on User Interface Software and Technology - 2012", Cambridge, MA, United States, Proceedings of the 25th Symposium on User Interface Software and Technology, ACM, October 2012, p. 261-270 [*DOI :* 10.1145/2380116.2380150], https://hal.inria.fr/hal-00721574

[78] C. PINDAT, E. PIETRIGA, O. CHAPUIS, C. PUECH.*Drilling into complex 3D models with gimlenses*, in "Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology", VRST '13, ACM, 2013, p. 223–230, http://doi.acm.org/10.1145/2503713.2503714

[79] N. SHADBOLT, T. BERNERS-LEE, W. HALL.*The Semantic Web Revisited*, in "IEEE Intelligent Systems", 2006, vol. 21, n^{o} 3, p. 96-101, http://doi.ieeecomputersociety.org/10.1109/MIS.2006.62

[80] O. SHAER, E. HORNECKER.*Tangible User Interfaces: Past, Present, and Future Directions*, in "Found. Trends Hum.-Comput. Interact.", January 2010, vol. 3, n^{o} 1–2, p. 1–137, http://dx.doi.org/10.1561/1100000026

[81] B. SHNEIDERMAN, C. PLAISANT.*Designing the User Interface, 4th Edition*, Pearson Addison Wesley, 2005

[82] D. SPELMEZAN, C. APPERT, O. CHAPUIS, E. PIETRIGA.*Controlling widgets with one power-up button*, in "Proceedings of the 26th annual ACM symposium on User interface software and technology", UIST '13, ACM, 2013, p. 71–74, http://doi.acm.org/10.1145/2501988.2502025

[83] D. SPELMEZAN, C. APPERT, O. CHAPUIS, E. PIETRIGA.*Side pressure for bidirectional navigation on small devices*, in "Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services", MobileHCI '13, ACM, 2013, p. 11–20, http://doi.acm.org/10.1145/2493190.2493199

[84] G. TOUYA, J.-F. GIRRES.*ScaleMaster 2.0: a ScaleMaster extension to monitor automatic multi-scales generalizations*, in "Cartography and Geographic Information Science", 2013, vol. 40, n^{o} 3, p. 192–200, http://dx.doi.org/10.1080/15230406.2013.809233

[85] G. TOUYA.*Social Welfare to Assess the Global Legibility of a Generalized Map*, in "International Conference on Geographic Information", Springer, 2012, p. 198–211, http://dx.doi.org/10.1007/978-3-642-33024-7_15

[86] T. TSANDILAS, C. APPERT, A. BEZERIANOS, D. BONNET.*Coordination of Tilt and Touch in One- and Two-Handed Use*, in "Proceedings of the 32nd international conference on Human factors in computing systems", CHI '14, ACM, 2014, p. 2001-2004, http://doi.acm.org/10.1145/2556288.2557088

[87] G. TUMMARELLO, R. DELBRU, E. OREN.*Sindice.com: Weaving the Open Linked Data*, in "Proc. of the 6th Int. Semantic Web Conf. (ISWC2007)", Springer-Verlag, 2007, p. 552–565

[88] C. WARE.*Information visualization: perception for design*, Elsevier, 2012

[89] S. ZHAI, P. O. KRISTENSSON, C. APPERT, T. H. ANDERSEN, X. CAO.*Foundational Issues in Touch-Screen Stroke Gesture Design - An Integrative Review*, in "Foundations and Trends in Human-Computer Interaction", December 2012, vol. 5, n^{o} 2, p. 97-205 [*DOI :* 10.1561/1100000012], https://hal.inria.fr/hal-00765046

[90] R. A. DE ALMEIDA, C. PILLIAS, E. PIETRIGA, P. CUBAUD.*Looking behind bezels: french windows for wall displays*, in "Proceedings of the International Working Conference on Advanced Visual Interfaces", AVI '12, ACM, 2012, p. 124–131, http://doi.acm.org/10.1145/2254556.2254581

Team INFINE-POST

INFormation NEtworks

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

RESEARCH CENTER
Saclay - Île-de-France

THEME
Networks and Telecommunications

Table of contents

Team INFINE-POST

Creation of the Team: 2018 January 01

Keywords:

Computer Science and Digital Science:
A1.2.3. - Routing
A1.2.5. - Internet of things
A1.2.6. - Sensor networks
A1.2.9. - Social Networks
A2.6.1. - Operating systems
A3.3.2. - Data mining
A3.4.1. - Supervised learning
A3.4.2. - Unsupervised learning
A3.5.1. - Analysis of large graphs

Other Research Topics and Application Domains:
B4.4. - Energy delivery
B4.4.1. - Smart grids
B6.3.2. - Network protocols
B6.4. - Internet of things
B6.6. - Embedded systems
B8.1.2. - Sensor networks for smart buildings
B8.2. - Connected city
B9.5.1. - Computer science
B9.6.1. - Psychology
B9.6.5. - Sociology

1. Team, Visitors, External Collaborators

Research Scientists
Aline Carneiro Viana [Team leader, Inria, Researcher, HDR]
Cedric Adjih [Inria, Researcher]
Emmanuel Baccelli [Inria, Researcher, HDR]

Technical Staff
Francisco Acosta [Inria, until Nov 2018]
Benoit Formet [Inria]
Raul Fuentes Samaniego [Inria, until Jun 2018]
Duc Tuyen Ta [Univ Paris-Sud, from Jul 2018]

PhD Students
Licia Amichi [Inria, from Oct 2018]
Rafael Costa [CAPES, from Jun 2017, co-tutelle]
Hirah Malik [Inria, from Nov 2017]
Iman Hmedoush [Inria, from Oct 2018, co-advised with EVA]
Douglas Teixeira [CAPES, from Apr 2018, co-tutelle]
Lucas Santos de Oliveira [EMBRACE, from May 2018]

Joao Batista Borges [EMBRACE, from Oct 2018]
Loic Dauphin [Inria, until Aug 2018]
Roni Shigueta [Univ. of Parana, until Jul 2018, co-tutelle]
Guangshuo Chen [Inria, until Apr 2018]

Visiting Scientists

Vinicius Braga [EMBRACE, Mar 2018]
Kléber Cardoso [EMBRACE, Mar 2018]
Sand Luz Correa [EMBRACE, Mar 2018]
Anelise Fonseca [EMBRACE, Feb 2018]
Mauro Fonseca [EMBRACE, Feb 2018]
Parikshit Hegde [Nokia, from May 2018 until Jul 2018]
Akshit Kumar [Nokia, from May 2018 until Jul 2018]
Julinda Stefa [Inria, from Jul 2018 until Sep 2018]

Administrative Assistant

Laurence Fontana [Inria]

2. Overall Objectives

2.1. Overall Objectives

Our general goal is to develop distributed mechanisms for optimizing the operation of networks both at the logical and physical levels. Despite the massive increases in transmission capacity of the last few years, one has every reason to believe that networks will remain durably congested, driven among other factors by the steadily increasing demand for video content, and the forecasted additional traffic due to machine-to-machine (M2M) communications. There is thus a strong need for smart protocols which transport requested information at the cheapest possible cost on the network.

INFINE team is engaged in research along two main themes: *Human-centric Networking*, and *Internet of Things*. These research activities encompass both theoretical research (on elaboration of models, algorithms, protocols and formal characterization of their performance), and applied research (to validate and/or experiment the proposed concepts against real networking scenarios). INFINE fits in the theme "Networks and Telecommunications" of the research field "Networks, Systems and Services, Distributed Computing" at Inria.

2.2. New challenging demands

Nowadays, the physical context in which we access communication networks has drastically changed. While in the past, Internet was mostly accessed through fixed desktop computers, users are now mobile about 52.2% of their time online [0]. In addition, while communicating machines used to be sparse and wired, with the advent of the Internet of Things we now evolve in a dense, interconnected environment of heterogeneous devices communicating via wireless and/or via wires.

This new context of Internet uses challenges several aspects of currently deployed networks. Some aspects pertain to the physical architecture of the Internet. In particular, at the core of the Internet, a drastic increase in volume of data traffic is anticipated due to the emergence of new applications, generalization of cloud services, or the advent of the Internet of Things (IoT) and Machine-to-Machine (M2M) communications. On the other hand, at the edge of the Internet, user mobility and today's pervasiveness of computing devices with increasingly higher capabilities (i.e., processing, storage, sensing) have a fundamental impact on the adequacy of algorithms and communication mechanisms.

[0] https://www.statista.com/statistics/241462/global-mobile-phone-website-traffic-share/

Other aspects concern the logical architecture of the network. For instance, currently deployed protocols at layers above IP must now carry massive publish-subscribe traffic, be human-aware, and support delay tolerant communications and paradigms for which they were not initially designed. While considering physical and logical aspects of networks, the INFINE team will pursue research activities combining theoretical and experimental approaches.

3. Research Program

3.1. Fundaments

Our general goal is to develop distributed mechanisms for optimizing the operation of networks both at the mentioned logical and physical levels of the architecture. Taking an information- and human-centric perspective, we envision networks as means to convey relevant information to users, while adapting to customary practices (in terms of context, interests, or content demands) of such users.

We now detail further our agenda along two main specific axes, namely Human-Centric Networking and Internet of Things, bearing in mind that we will develop generic solutions relevant to several of these axes wherever possible.

Finally, at the end of this section, we also detail our activities related to standization and open experimental IoT Platforms.

3.1.1. Human-Centric Networking

Cellular and wireless data networks are increasingly relied upon to provide users with Internet access on devices such as smartphones, laptops or tablets. In particular, the proliferation of handheld devices equipped with multiple advanced capabilities (e.g., significant CPU and memory capacities, cameras, voice to text, text to voice, GPS, sensors, wireless communication) has catalyzed a fundamental change in the way people are connected, communicate, generate and exchange data. In this evolving network environment, users' social relations, opportunistic resource availability, and proximity between users' devices are significantly shaping the use and design of future networking protocols.

One consequence of these changes is that mobile data traffic has recently experienced a staggering growth in volume: Cisco has recently foreseen that the mobile data traffic will increase 18-fold within 2016, in front of a mere 9-fold increase in connection speeds. Hence, one can observe today that the inherently centralized and terminal-centric communication paradigm of currently deployed cellular networks cannot cope with the increased traffic demand generated by smartphone users. This mismatch is likely to last because (1) forecasted mobile data traffic demand outgrows the capabilities of planned cellular technological advances such as 4G or LTE, and (2) there is strong skepticism about possible further improvements brought by 5G technology.

Congestion at the Internet's edge is thus here to stay. Solutions to this problem are either to densify infrastructure, or to offload to alternate networks. Densifying infrastructure (with Femtocells for instance) is expensive. Carriers therefore consider other solutions, such as simultaneously leveraging Wi-Fi access points and hot spots. However, the projected increase of mobile data traffic demand pushes towards additional complementary offloading methods. Novel mechanisms are thus needed, which must fit both the new context that Internet users experience now, and their forecasted demands.

In this realm, we focus on new approaches leveraging ultra-distributed, user-centric approaches over IP. One approach that is considered in the research community is to leverage spontaneous wireless networks to offload infrastructure-based cellular networks. For example, the delay-tolerant nature of some of the data traffic can be used in conjunction with sequences of opportunistic encounters between users to deliver such data to its destination, without infrastructure access point relaying. However, the full capacity and the achievable information propagation speed in such networks are still barely understood, and in particular, there is a need to refine the characterization of user behaviour and social interaction. Our recent work on leveraging user mobility patterns, contact and inter-contact patterns, content demand patterns will constitute a starting point.

3.1.2. Internet of Things at the Edge

The IoT is indeed expected to massively use this networking paradigm to gradually connect billions of new devices to the Internet, and drastically increase communication without human source or destination – to the point where the amount of such communications will dwarf communications involving humans. Over the last decade, we are witnessing an increasing variety in IoT technologies starting from IoT communication technologies. The main reason for this is the growing diversity of strong requirements (i.e., in terms of bandwidth, latency, energy savings, etc) coming from new varieties of IoT use-cases, which now go far beyond the original wireless sensor networking domain. Besides, such variety of new IoT devices still inherits from the extremely limited capabilities of traditional embedded systems and sensor networks, such as requirements in terms of low power usage, low memory, with today a greater emphasis on interoperability needs.

Large scale user environment automation require communication protocols optimized to efficiently leverage the heterogeneous and unreliable wireless vicinity (the scope of which may vary according to the application). In fact, extreme constraints in terms of cost, CPU, battery and memory capacities are typically experienced on a substantial fraction of IoT devices. We expect that such constraints will not vanish any time soon for two reasons. On one hand the progress made over the last decade concerning the cost/performance ratio for such small devices is quite disappointing. On the other hand, the ultimate goal of the IoT is ubiquitous Internet connectivity between devices as tiny as dust particles. These constraints actually require to redesign not only the network protocol stack running on these devices, but also the software platform powering these machines.

In this context, we will aim at contributing to the design of novel network protocols and software platforms optimized to fit these constraints while remaining compatible with legacy Internet. Our recent work on large IoT testbeds such as FIT and on software platforms such as the RIOT operating system will serve as a starting point to design and conduct large scale experiments that are expected to provide both fruitful feedback to our theoretical analysis, and validation of the protocols we propose in the realm of our standardisation activities. A recent example of this methodology is the publication of RFC6997, the specification of a routing protocol for sensor networks we proposed, which was standardized a few months ago. An example of technique which we plan to explore in this field is the use of network coding. Network coding is ideally suited to such spontaneous wireless networks for increasing communication reliability (while minimizing the traffic load) ; exploiting it fully requires keeping track of the information flows, the central topic of this project proposal. In this domain, we will use as starting point our latest work on practical network coding broadcast.

Information centric networking paradigms emerge to decouple data name and location and organizes pervasive content caching and nearest replica routing, promising performance gains in terms of native multi-homing optimization, content access time and network load, at the price of more complex, more voluminous and volatile state management in routers. In this context, we investigate generic network protocols, that provide a significantly advantageous tradeoff between performance gains and required router state complexity increase on low-end IoT networks.

3.1.3. Open Experimental IoT Platforms

One necessity for research in the domain of IoT is to establish and improve IoT hardware platforms and testbeds, that integrate representative scenarios (such as Smart Energy, Home Automation etc.) and follow the evolution of technology, including radio technologies, and associated experimentation tools. For that, we plan to build upon the IoT-LAB federated testbeds, that we have participated in designing and deploying recently. We plan to further develop IoT-LAB with more heterogeneous, up-to-date IoT hardware and radios that will provide a usable and realistic experimentation environment. The goal is to provide a tool that enables testing a validation of upcoming software platforms and network stacks targeting concrete IoT deployments.

In parallel, on the software side, IoT hardware available so far made it uneasy for developers to build apps that run across heterogeneous hardware platforms. For instance Linux does not scale down to small, energy-constrained devices, while microcontroller-based OS alternatives were so far rudimentary and yield a steep learning curve and lengthy development life-cycles because they do not support standard programming and debugging tools. As a result, another necessity for research in this domain is to allow the emergence of it more powerful, unifying IOT software platforms, to bridge this gap. For that, we plan to build upon RIOT, a new open source software platform which provides a portable, Linux-like API for heterogeneous IoT hardware. We plan to continue to develop the systems and network stacks aspects of RIOT, within the open source developer community currently emerging around RIOT, which we co-founded together with Freie Universitaet Berlin. The key challenge is to improve usability and add functionalities, while maintaining architectural consistency and a small enough memory footprint. The goal is to provide an IoT software platform that can be used like Linux is used for less constrained machines, both (i) in the context of research and/or teaching, as well as (ii) in industrial contexts. Of course, we plan to use it ourselves for our own experimental research activities in the domain of IoT e.g., as an API to implement novel network protocols running on IoT hardware, to be tested and validated on IoT-LAB testbeds.

3.1.4. Standardization of Architectures and Efficient Protocols for Internet of Things

As described before, and by definition, the Internet of Things will integrate not only a massive number of homogeneous devices (e.g., networks of wireless sensors), but also heterogeneous devices using various communication technologies. Most devices will be very constrained resources (memory resources, computational resources, energy). Communicating with (and amongst) such devices is a key challenge that we will focus on. The ability to communicate efficiently, to communicate reliably, or even just to be able to communicate at all, is non-trivial in many IoT scenarios: in this respect, we intend to develop innovative protocols, while following and contributing to standardization in this area. We will focus and base most of our work on standards developed in the context of the IETF, in working groups such as 6lo, CORE, LWIG etc., as well as IRTF research groups such as NWCRG on network coding and ICNRG on Information Centric Networking. We note however that this task goes far beyond protocol design: recently, radical rearchitecturing of the networks with new paradigms such as Information Centric Networking, ICN, (or even in wired networks, software-defined networks), have opened exciting new avenues. One of our direction of research will be to explore these content-centric approaches, and other novel architectures, in the context of IoT.

4. Highlights of the Year

4.1. Highlights of the Year

4.1.1. Awards

Together with his co-authors, Emmanuel Baccelli was awarded the **best demo award** at the 3rd Cloudification of the Internet of Things Conference, in Paris, July 2018, for the demo on **Orchestration of IoT Device and Business Workflow Engine on Cloud** (collaboration with S. Kikuchi, I. Thomas, O. Jallouli, J. Dörr, A. Morgenstern, and K. Schleiser).

RIOT Summit 2018

We successfully organized in September 2018 the thrid RIOT Summit, in Amsterdam. The RIOT Summit 2018 gathered 100+ enthusiastic industrial participants, makers and academics involved in RIOT. Relevant partners such as Ericsson, HERE Technologies, CodeCoup, Wolf SSL, as well as a number of SMEs and startups from various places in Europe gave talks on aspects of IoT communication, use cases IoT hardware, IoT open source community aspects and concepts for future IoT software and networks, as well as hands-on sessions and tutorials. See: http://summit.riot-os.org.

Associated team - EMBRACE

2018 was the second year of the EMBRACE Associated team. The EMBRACE (lEveraging huMan Behavior for Resource AlloCation and services orchestration modEls) team is composed by members of the INFINE and by three Brazilian teams from three different Brazilian Universities. The EMBRACE project addresses the topic of designing efficient solutions for 5G networks taking into account human behavior, uncertainty, and heterogeneity of networking resources.

More information is available here: https://team.inria.fr/embrace/

5. New Results

5.1. IoT Scripting Over-The-Air

Participants: Emmanuel Baccelli, Francisco Acosta.

A large part of the Internet of Things (IoT) will consist of interconnecting low-end devices, whose characteristics include very small memory capacity (a few kBytes) and limited energy consumption (1000 times less than a RaspberryPi). IoT use-cases require the orchestration of different pieces of logic running concurrently on low-end IoT devices and elsewhere on the network (e.g. in the cloud) and communicating with one another. In a number of use-cases, the logic that needs to run on low-end IoT devices is not known upfront, before deploying the device(s). For instance, some part of the logic (e.g. pre-processing of some data) may need to be transferred on demand, from the cloud to the device, for privacy or performance reasons. Another example is the fine-tuning of some parameters of the logic running on some device, which can only be done after the deployment (e.g. the sensitivity of a distributed alarm system on-site). In such context, this paper presents a generic approach to host, run and update IoT application logic on heterogeneous low-end devices, using over-the-air scripting and small containers. Based on RIOT and Javascript, we provide a proof- of-concept implementation of this approach for a building automation IoT scenario, as well as a preliminary evaluation of this implementation running on common off-the-shelf low-end IoT hardware. Our evaluation shows the prototype runs on common off-the-shelf low-end IoT hardware with as little as 32kB of memory. Recent prior work in this domain also proposed Actinium, an approach using small, distributed runtime containers on computers proxying for low-end IoT devices, accessible as Web resources, and hosting JavaScript logic. Compared to Actinium, we eliminate the need for Web resource proxying, as runtime containers are running directly on the low-end IoT devices.

This work was published and presented at the IEEE Percom 2018 conference as "Scripting Over-The-Air: Towards Containers on Low-end Devices in the Internet of Things".

5.2. Information-centric IoT Robotics

Participants: Loic Dauphin, Cedric Adjih, Emmanuel Baccelli.

As IoT emerges, minibots (miniature robots) have appeared on the market. A large community emerged, designing do-it- yourself minibots, and cheap, re-programmable minibots with communication capabilities are now available. For in- stance, small wheeled robots such as the Zooid are based on a small microcontroller (8kB RAM, 64kB ROM) and communicating with a low-power radio in the 2.4 GHz ISM band. Other examples are cheap drones such as the Cheerson CX-10, which has similar hardware characteristics, and which costs under 15$. Simple robotic arms and legged robots are also available, such as the MetaBot. A current trend bases software embedded in minibots on open source frameworks. The Robot Operating System (ROS) is a software framework for robot application development which has become a de facto standard for most areas in robotics. Other open source robotics frameworks include software suite tailored for drones, some of which provide compatibility with ROS. In fact, we observe that minibots have a number of charac- teristics in common with low-end devices found in the Internet of Things (IoT). Compared to low-end IoT devices, minibots are based on similar hardware and their software follows similar trends. For instance, an IoT-enabled actuator based on a System-on-Chip (SoC) embarking a small microcontroller, and a radio communicating with a remote server, is very similar to a simple radio-controlled robot. Low-end IoT devices use similar radio modules, and software embedded in IoT devices is more and more based on a variety of open source, lightweight operating systems such as RIOT, FreeRTOS and NuttX, among others. Similarly, as for IoT embedded systems, the network com- ponent of minibots represents by itself in important part of the software (in terms of features, code/memory size, and performance). In fact, a wide variety of radio modules and communication protocols are used on minibots. The protocols used by micro-robots for (internal or external) communication range from direct motor control (pulse width modulation PWM, pulse position modulation PPM, or PCM), to serial/bus protocols, and high level protocols such as Real-time Publish- Subscribe Protocol (RTPS). In this work we thus to explored the potential of bundling open source robotics software frameworks with IoT software and network architectures, to program and control minibots. To do so, we extend our recent work by designing ROS-ready technology for a minibot based on RIOT and ROS2. We focus primarily on software and networking aspects, targeting ultra-lightweight robots based on a reprogrammable SoC with a microcontroller running at approximately 50 MHz, with 10kB RAM, 100kB Flash, and a low-power radio. Using an information-centric networking paradigm extending NDN, we design and implement the communication primitives required by RIOT-ROS2. Our prototype is able to maintain full compatibility between ROS nodes running on the minibot(s) and ROS nodes running elsewhere on the network without the use of a bridge. We show that RIOT-ROS2 fits on low-end robotics hardware such as a System-on-Chip with an ARM Cortex-M0+ microcontroller. On the software and network performance evaluation side, we illustrate that the latency incurred with our ICN approach is completely acceptable for minibot control, even on constrained radio, based on micro-benchmarks.

This work was published and presented at the IEEE PEMWN 2018 conference as "RIOT-ROS2: Low-Cost Robots in IoT Controlled via Information-Centric Networking".

5.3. Human Mobility completion of Sparse Call Detail Records

Participants: Guangshuo Chen, Aline Carneiro Viana, Marco Fiore [CNR - IEIIT (Italy)], Carlos Sarraute [Grandata Labs].

Mobile phone data are a popular source of positioning information in many recent studies that have largely improved our understanding of human mobility. These data consist of time-stamped and geo-referenced communication events recorded by network operators, on a per-subscriber basis. They allow for unprecedented tracking of populations of millions of individuals over long time periods that span months. Nevertheless, due to the uneven processes that govern mobile communications, the sampling of user locations provided by mobile phone data tends to be sparse and irregular in time, leading to substantial gaps in the resulting trajectory information. In this work, we illustrate the severity of the problem through an empirical study of a large-scale Call Detail Records (CDR) dataset. We then propose two novel and effective techniques to reduce temporal sparsity in CDR that outperform existing ones. the fist technique performs completion (1) at nightime by identifying temporal home boundary and (2) at daytime by inferring temporal boundaries of users, i.e., the time span of the cell position associated with each communication activity. The second technique, named Context-enhanced Trajectory Reconstruction, complete individual CDR-based trajectories that hinges on tensor factorization as a core method by leveraging regularity in human movement patterns. Our approach lets us revisit seminal works in the light of complete mobility data, unveiling potential biases that incomplete trajectories obtained from legacy CDR induce on key results about human mobility laws, trajectory uniqueness, and movement predictability.

These works have been published as invited papers at the ACM CHANTS 2016 workshop (in conjunction with ACM MobiCom 2016), at the IEEE DAWM workshop (in conjunction with IEEE Percom 2017) and at Computer Communication Elsevier journal in 2018. Another journal version (also registered as TR: hal-01675570) is in revision at the EPJ Data Science Journal.

5.4. Adaptive sampling frequency of human mobility

Participants: Panagiota Katsikouli, Aline Carneiro Viana, Marco Fiore [CNR - IEIIT (Italy)], Diego Madariaga.

The problem we address here is the design of a location sampling system for smartphones and handheld devices that reduces the energy consumed by the continuous activation of the GPS, it reduces the space required to store recorded locations, while reliably capturing the movements of the tracked user. The applications here are related to a number of fields relevant to ubiquitous computing, such as energy-efficient mobile computing, location-based service operations, active probing of subscribers' positions in mobile networks and trajectory data compression.

To this end, we propose an adaptive sampling system without the use of any assisting sensors for the activation of GPS, such as accelerometer, or GSM information. Our system captures the mobility of a user with high accuracy and reliably adjusts the sampling frequency depending on the user's movement. During high mobility, our system densely samples the locations of the tracked user, but at a rate at most the usual rate found today in most applications (e.g., 1 sample per minute). During low mobility, we sample sparsely at much lower rate than usual. As a result, the recorded trace contains much less samples than it would contain if we sampled with the fixed pre-defined sampling rate, requiring less storage space and less energy to activate the GPS.

Our first quest for a response led to the discovery of *(i)* seemingly universal spectral properties of human mobility, and *(ii)* a linear scaling law of the localization error with respect to the sampling interval. Our findings were based on the analysis of fine-grained GPS trajectories of 119 users worldwide. This work was published at the IEEE Globecom 2017 international conference.

We have improved the published sampling approach by incorporating human behavioral features at the sampling decisions to make it more adaptive. This is an on-going work with Panagiota Katsikouli, who spent 5 months in our team working as an internship and is currently doing a Post-Doc at the AGORA Inria team, and Diego Madariaga who spent 3 months in our team working as an internship and is going to start a PhD in co-tutelle with Aline C. Viana. Diego has implemented an Android application to sample mobility data of users according to our adaptive system described here above. The application is currently under deployment and 8 volunteers are running it in their smartphones. The collected data will allow us validating the correctness and performance of our adaptive sampling system. A patent discussion is also on-going with Inria, currently performing a marked/business study.

5.5. Inference of human personality from mobile phones datasets

Participants: Adriano Di Luzio, Aline Carneiro Viana, Julinda Stefa, Katia Jaffres-Runser [INPT-ENSEEIHT - IRIT (Toulouse University)], Alessandro Mei [Sapienza University (Italy) - Dept. of Computer Science].

Related to human behavioral studies, personality prediction research has enjoyed a strong resurgence over the past decade. Due to the recognition that personality is predictive of a wide range of behavioral and social outcomes, the human migration to the digital environment renders also possible to base prediction of individual personality traits on digital records (i.e., datasets) mirroring human behaviors. In psychology, one of the most commonly used personality model is the Big5, based on five crucial traits and commonly abbreviated as OCEAN: Openness (O), Conscientiousness (C), Extroversion (E), Agreeableness (A), and Neuroticism (N). They are relatively stable over time, differ across individuals, and, most importantly, guide our emotions and our reactions to life circumstances. It is so for social and work situations, and even for things as simple as the way we use our smartphone. For instance, a person that is curious and open to new experiences will tend to look continuously for new places to visit and thrills to experience.

This work brings the deepest investigation in the literature on the prediction of human personality (*i.e.,* captured by the Big5 traits) from smartphone data describing daily routines and habits of individuals. We take a ground-breaking step in (i) deeply capturing human habits in terms of movements, visits, wireless connectivity as well as some routinary actions from a crowdsourced mobility dataset and in (ii) better understanding the relationship between personality traits and individual behavior. We do so by leveraging a dataset collecting very detailed routines of individuals originating from different countries located in 2 different continents, who answered the Big Five Inventory and allowed continuous collection of data from their smartphones for research purposes for 3 years. We use this dataset to engineer a set of human-adapted features that capture three aspects of human behavior: Temporal Mobility (e.g. time at home/work or commuting), Spatial Mobility (e.g. number of most frequent places, maximum distance from home), and the Context of Use (battery charging habits, wireless hotspots availabilities). Then, we use the features that have a statistically significant correlation with the OCEAN traits to predict the personality of a test-set portion of our dataset through cross validation.

Our results attest an accurate prediction of users' personality traits when a 5-level granularity is used per trait. This brings a much higher precision to our predicted results, when compared to the usual 3-level literature granularity. In addition, our prediction methodology carefully takes advantage of engineered features that (1) are more human-adapted and consequently, allow better capturing individuals' habits in terms of movements, visits, connectivity, context, as well as actions (note that contrarily to the literature, neither calls behavior nor data content is leveraged in our analysis), and (2) are designed having in mind the differences and particularities among the Big5 traits of personality. Thus, this work has the potential to impact the way we characterise unique behaviors of individuals as well as quantify how human personality influences lives and actions. Our results show (1) a significant correlation of most of the traits with a small set of mobility-related features and (2) that we are able to predict the individuals' Big5 traits with considerable accuracy (e.g., prediction of the 5 levels of Openness trait shows an F1 score of 0.77), which is significantly outperforming a benchmark approach, when only considering a set of only 3 of our human-adapted features. Finally, we discuss the ethical concerns of our work, its privacy implications, and ways to tradeoff privacy and benefits.

This is an on-going work with Adriano di Luzio, who spent 4 months in our team working as an internship, Julinda Stefa, an invited research visitor at Infine, and two other researchers: Katia Jaffres-Runser and Alessandro Mei. A paper describing this work is under submission at ACM Mobihoc 2018, but a technical report is also registered under the name hal-01954733.

5.6. Data offloading decision via mobile crowdsensing

Participants: Emanuel Lima, Aline Carneiro Viana, Ana Aguiar [FEUP (Portugal) - Dept. of Electrical and Computer Engineering], Paulo Carvalho [FEUP (Portugal) - Dept. of Electrical and Computer Engineering].

According to Cisco forecasts [0], mobile data traffic will grow at a compound annual growth rate of 47 % from 2016 to 2021 with smartphones surpassing four-fifths of mobile data traffic. It is known that mobile network operators are struggling to keep up with such traffic demand, and part of the solution is to offload communications to WiFi networks. Mobile data offloading systems can assist mobile devices in the decision making of when and what to offload to WiFi networks. However, due to the limited coverage of a WiFi AP, the expected offloading performance of such a system is linked with the users mobility. Unveiling and understanding human mobility patterns is a crucial issue in supporting decisions and prediction activities for mobile data offloading.

Several studies on the analysis of human mobility patterns have been carried out focusing on the identification and characterization of important locations in users' life in general. We intend to extend these works by studying human mobility from the perspective of mobile data offloading. This brings two major differences compared to the related work. First, high temporal resolution of positioning datasets is needed. In the majority of the related work, important locations have a temporal dimension representing the time spent by a user in that location, which confers its degree of importance. This time is usually in the order of several minutes which is suitable for the case of detecting important locations but not for a mobile data offloading scenario. Here, according to the amount of data traffic that needs to be offloaded, locations with a visiting temporal resolution of few seconds may be enough for data offloading. Thus, we expect to discover additional offloading opportunities, which were not visible with a coarser temporal resolution. Second, while important locations are usually limited in size, offloading locations can have any arbitrary shape and size.

In this work, offloading regions are defined as spatially aggregated locations where users have mobility suitable to offload. The main contribution of this work are: (a) the identification of offloading regions on an individual basis through unsupervised learning; (b) the characterization of these regions in terms of availability, sojourn, and transition time based on their relevance; (c) the study of the impact of the users mobility on the design of mobile offloading systems. This work was published at ACM CHANTS 2018.

We now working on the extension of this work, which will incorporate the mobility prediction of the users. Such prediction is essential to the design of the decision offloading strategy. Such strategy will be used to allow a mobile phone of a user deciding if offload or not her traffic, i.e., when, where (in which offloading region) and how (if the traffic will be offloadied to one or more Access Points). This is an on-going work with the the PhD Emanuel Lima, who spent 4 months as an intern in our team, and his advisors.

[0] https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html

5.7. Infering friends in the crowd in Device-to-Device communication

Participants: Rafael Costa, Aline Carneiro Viana, Leobino Sampaio [UFBA (Brazil) - Institute of Mathematics], Artur Ziviani [National Laboratory for Scientific Computing (Brazil)].

The next generation of mobile phone networks (5G) will have to deal with spectrum bottleneck and other major challenges to serve more users with high-demanding requirements. Among those are higher scalability and data rates, lower latencies and energy consumption plus reliable ubiquitous connectivity. Thus, there is a need for a better spectrum reuse and data offloading in cellular networks while meeting user expectations. According to literature, one of the 10 key enabling technologies for 5G is device-to-device (D2D) communications, an approach based on direct user involvement. Nowadays, mobile devices are attached to human daily life activities, and therefore communication architectures using context and human behavior information are promising for the future. User-centric communication arose as an alternative to increase capillarity and to offload data traffic in cellular networks through opportunistic connections among users. Although having the user as main concern, solutions in the user-centric communication/networking area still do not see the user as an individual, but as a network active element. Hence, these solutions tend to only consider user features that can be measured from the network point of view, ignoring the ones that are intrinsic from human activity (e.g., daily routines, personality traits, etc). In this work, we plan to investigate how human-aspects and behavior can be useful to leverage future device-to-device communication.

This is the PhD thesis subject of Rafael Costa, aiming the design of a methodology to select next-hops in a D2D communication that will be human-aware: i.e., that will consider not only available physical resources at the mobile device of a wireless neighbor, her mobility features and restrictions but also any information allowing to infer how much sharing willing she is. A tutorial paper is under submission to a journal (a TR is in hal-01675445) and a 4h-tutorial was presented at the SBRC 2018 conference [0] (the biggest conference on Computer and Network Science in Brazil).

The next step is then the design of forwarding strategies for data offloading through Device-to-Device (D2D) communication, transforming mobile phone neighbors in service providers. The selection of next hops based on mobility behavior, resource capability as well as collaboration constitute the novelty we plan to exploit.

5.8. Urban Computing Leveraging Location-Based Social Network Data: a Survey

Participants: Thiago H. Silva [UTFPR (Brazil) - Dept. of Computer Science], Aline Carneiro Viana, Antonio Loureiro.

Urban computing is an interdisciplinary area in which urban issues are studied using state-of-the-art computing technologies. This area is at the intersection of a variety of disciplines: sociology, urban planning, civil engineering, computer science, and economics, to name a few. More than half of the world's population today live in cities and, consequently, there is enormous pressure on providing the proper infrastructure to cities, such as transport, housing, water, and energy. To understand and partly tackle these issues, urban computing combines various data sources such as those coming from Internet of Things (IoT) devices; statistical data about cities and its population (e.g., the Census); and data from Location-Based Social Networks (LBSN), sometimes also termed as location-based social media. One fundamental difference between data from LBSNs and data from other sources is that the former offers unprecedented geographic and temporal resolutions: it reflects individual user actions (fine-grained temporal resolution) at the scale of entire world-class cities (global geographic resolution).

Urban computing with LBSN data has its particularities. For instance, users who share data in Foursquare, a popular LBSN, usually have the goal of showing to their friends where they are while also providing personalized recommendations of places they visit. Nevertheless, when correctly analyzed for knowledge extraction, this data can be used to better understand city dynamics and related social, economic, and cultural aspects. To achieve this purpose, new approaches and techniques are commonly needed to explore that data properly.

[0] http://www.sbrc2018.ufscar.br/minicurso-1-mc-1/

In order to better study such needs, we have published at ACM Computing Survey Journal (the ACM journal with highest impact factor) a survey that provides an extensive discussion of the related literature, focusing on major findings and applications. Although its richness concerning knowledge provision, LBSN data presents several challenges, requiring extra attention to its manipulation and usability, which drives future research opportunities in the field of urban computing using LBSN data. Our work is complementary to two existing surveys in the area of urban computing (i.e., by Jiang et al. and by Zheng et al.) since they only mention briefly few studies that explore LBSN data, neglecting key challenges that revolve around LBSNs. We hope that taken together, our effort and these existing ones, provide a broad perspective of urban computing studies and its development through the lens of different data-driven approaches.

5.9. Identifying how places impact each other by means of user mobility

Participants: Lucas Santos, Pedro Olmo [UFMG (Brazil) - Dept. of Computer Science], Aline Carneiro Viana.

The way in which city neighborhoods become popular and how people trajectory impacts the number of visitation is a fundamental area of study in traditional urban studies literature. Many works address this problem by means of user mobility prediction and POI recommendation. In a different approach, other works address the human mobility in terms of social influence which refers to the case when individuals change their behaviors persuaded by others. Nevertheless, fewer works measure influence of POI based on human mobility data.

Different from previous literature, in this work, we are interested in understanding how the neighborhood POI affect each other by means of human mobility using location-based social networks (LBSNs) data source. In other words, how important is this POI for its neighborhood? We proposed thus a framework to measure POI influence by means of LBSN data. First, we modeled the problem using mobility graph approach where each POI is a node and the transitions of users among POI is a weighted vertex. Also, we treat the users' check-in records among POI as a measure of uncertainty, and their strength can be measured by entropy, which enabled to measure direct influence. Second, using same graph, we propose another influence measure taking account the POI importance for its one-hop vicinity in terms of incoming human transition. In addition, this mobility graph can be viewed as a collaborative filtering. We use this collaborative filter for compute the G-causality and evaluate if the transitions among POI has a causal relation and consequently, the influence among POI. Moreover, to the best of our knowledge, we are the first study which investigated POI influence by means of human mobility using LBSN data source.

This work is being prepared for a submission to an international conference.

6. Bilateral Contracts and Grants with Industry

6.1. Bilateral Contracts with Industry

6.1.1. Fujitsu (RunMyProcess):

Participants: Emmanuel Baccelli, Francisco Acosta.

In 2018 we have worked with Fujitsu RIOT enhancements to demonstrate dynamic application software loading and execution on top of RIOT running on Arduino-like hardware, managed remotely from Fujitsu's RMP Cloud component. The results of this work were published in several conferences in 2018, and a prototype was demonstrated.

6.1.2. Thalès:

Participant: Cedric Adjih.

In 2018, studies were made with Thalès (TRT) on IoT systems.

6.1.3. GranData:

Participants: Guangshuo Chen, Adriano Di Luzio, Aline Carneiro Viana.

Since June 2014, we have a collaboration with GranData (http://grandata.com/), Buenos Aires, Argentina on traffic vs mobility modeling of smartphone users. GranData is a small company that integrates first-party and telco partner data to understand key market trends, to predict customer behavior, and to deliver business results. For the time being, the collaboration with Grandata has generated knowledge transfer. From both directions, (1) from myself to GranData, I have being transferring my knowledge in modeling and analysing human behavior in terms of mobility, encounters, and content demand, (2) from them to myself, they have advising me on issues related to machine learning and statistical methods to be used. It describes **an industrial partner's collaboration having the outcomes of our works impacting their products** (e.g., GranData data mining algorithms can be improved based on the better understanding on mobility and content consumption of mobile users) **or research/business decisions** (e.g., proved strong correlations between mobility and data traffic consumption can open new perspectives of services to telecom operators, i.e., clients of GranData).

Part of the thesis of Guangshuo Chen (endend April 2018) and of Eduardo Mucelli (ended in 2015) on data traffic analysis used telco traces provided by GranData.

7. Partnerships and Cooperations

7.1. National Initiatives

7.1.1. Equipex FIT:

Participants: Cedric Adjih, Emmanuel Baccelli, Alexandre Abadie [SED - Inria], Ichrak Amdouni [Ecole Nationale d'Ingénieurs de Sousse & CRISTAL].

Partners: Inria (Lille, Sophia-Antipolis, Grenoble), INSA, UPMC, Institut Telecom Paris, Institut Télécom Evry, LSIIT Strasbourg.

FIT (Future Internet of Things) aims to develop an experimental facility, a federated and competitive infrastructure with international visibility and a broad panel of customers. It provides this facility with a set of complementary components that enable experimentation on innovative services for academic and industrial users. The project gives french internet stakeholders a means to experiment on mobile wireless communications at the network and application layers thereby accelerating the design of advanced networking technologies for the future internet. FIT was one of 52 winning projects from the first wave of the French Ministry of Higher Education and Research's "Équipements d'Excellence" (Equipex) research grant program, in 2011.

One component of the FIT platform is the sets of IoT-LAB testbeds (see the IoT-LAB web site). These were motivated by the observation that the world is moving towards an "Internet of Things", in which most communication over networks will be between objects rather than people.

The Infine team is more specifically managing the FIT IoT-LAB site formerly at Rocquencourt, which recently moved to Saclay (on-going re-deployment), and is participating in the deployment of an additional IoT-lab testbed in Berlin (at Freie Universitaet Berlin).

7.1.2. ANR

The team has submitted three PRC ANR projects: two projects for the CE25, i.e., on "Réseaux de communication multi-usages, infrastructures de hautes performances, sciences et technologies logicielles" and one for the CE35, i.e., on "Révolution numérique : rapports au savoir et à la culture".

7.2. European Initiatives

7.2.1. H2020 Projects

.

7.2.2. AGILE (H2020 project)

Participants: Emmanuel Baccelli, Cedric Adjih.

Program: H2020 ICT-30-2015 Topic: Internet of Things and Platforms for Connected Smart Objects

Project acronym: AGILE

Project title: Adoptive Gateways for dIverse muLtiple Environments

Duration: 2015-2018

Coordinator: Emmanuel Baccelli

Other partners: Canonical (UK), Eclipse IoT Foundation (IE), Mobistar (BE), Libelium (ES), Startupbootcamp IoT (SP), CREATE-NET (IT), iMinds (BE), Atos (SP), Rulemotion (UK), Jolocom (DE), Passau University (DE), Sky-Watch (DN), BioAssist (GR), Graz Technical University (AT), Eurotech (IT), IoTango (US).

Abstract: The AGILE project is a 3-year H2020 project started in January 2016, which will deliver an integrated framework of open source tools and platforms that interoperate for enabling the delivery of adaptive, self-configurable and secure IoT elements (both software and hardware) that can be utilized in a variety of scenarios. Such tools target actors with heterogeneous skills, including entrepreneurs, researchers, and individuals, aiming to enable the realization of IoT applications respecting user privacy and data ownership.

7.3. International Initiatives

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

7.3.1.1. EMBRACE

Title: Leveraging Human Behavior and Uncertainty in 5G Networks to Build Robust Resource Allocation and Services Orchestration Models

International Partners (Institution - Laboratory - Researcher):

UTFPR (Brazil) - Departamento Acadêmico de Informática (DAINF) Curso de Pós-Graduação em Engenharia Elétrica e Informática Industrial (CPGEI) - Anelise Munaretto

UFG (Brazil) - Institute of Computational Mathematics and Scientific / Engineering Computing - Kleber Vieira Cardoso

UFMG (Brazil) - Dpt of Statistics - Antonio A. F. Loureiro

Start year: 2017

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

Abstract: EMBRACE propose une architecture novatrice pour gérer des ressources et des services opérationnels hétérogènes. EMBRACE se concentre sur les défis scientifiques liés des ensembles de données collectées dans le monde réel et décrivant le comportement du réseau des utilisateurs. En particulier, EMBRACE exploite la modélisation du comportement humain en termes de mobilité, de demande de contenu, d'intérêts communs et des interactions entre-utisateurs. En construisant des modèles d'allocation les ressources tenant compte de l'utilisateur, EMBRACE a pour objectif de diminuer l'incertitude et mieux cerner les profils humains dans les réseaux 5G. La communication D2D sera également utilisée comme service opérationnel pour gérer la croissance du trafic mobile en libérant des ressources des réseaux cellulaires, sans augmenter les coûts. La nouveauté de l'architecture réside dans les algorithmes conçus qui exploiteront les caractérisations tirés de l'analyse du comportement des utilisateurs, l'hétérogénéité du réseau, et de l'incertitude. L'évaluation par simulation et l'émulation sera également l'un des thèmes clés. Enfin, les équipes concernées (Inria Infine, UFMG, UFG, UTFPR) ont un long historique de coopération sur ces thèmes.

7.3.2. Inria International Partners

7.3.2.1. Declared Inria International Partners

1. Renewed IOTPUSH collaboration with Freie Universitaet Berlin around the long-term stay of Emmanuel Baccelli in Berlin, on research topics about the Internet of Things, RIOT and Information-Centric Networking.

7.3.2.2. Informal International Partners

1. On-going collaboration with Freie Universitaet Berlin and Hamburg University of Applied Science around RIOT.
2. Informal collaborations with UIUC and UMass.
3. Informal collaborations with ENSI Tunis and Sesame Tunis.
4. On-going strong collaboration with Sapienza University of Rome, Italy.
5. On-going strong collaboration with CNR Torino, Italy.
6. On-going collaboration with University of Porto, Portugal.
7. On-going collaboration with ENSAE/CNRS, France.
8. On-going collaboration with University of Edinburgh, UK.
9. On-going collaboration with Boston University, US.

7.3.3. Participation in Other International Programs

7.3.3.1. Indo-French project

The Inria teams Infine and Eva are part of the "D2D Communication for LTE Advanced Cellular Network" , a project funded by the Indo-French Centre for the Promotion of Advanced Research (CEFIPRA). With industrial partners, and also with Indian partners, this project is focusing on the evolution of cellular networks towards 5G: this includes exploration of device-to-device (D2D) communication, and more generally IoT communication in a cellular context. Research directions include efficient access for IoT devices (massive numbers of devices with low volume communication); combination of random access protocols/error coding/physical layer ; efficient neighbor discovery, ...

7.3.3.2. STIC AmSud MOTIf 2017

Participant: Aline Carneiro Viana.

Program: STIC AmSud

Project title: Mobile phone sensing of human dynamics in techno-social environment

Duration: 2017-2019

Coordinators: Marton Karsai (ENS/Inria) and Jussara M. Almeida (UFMG) and Alejo Salles (Univ. of Buenos Aires)

Abstract: Information and Communication Technology (ICT) is becoming increasingly social, as demonstrated by the multitude of emerging technologies and technology platforms that facilitate social interactions, taking place as communication via telephone, text message, email, online social networks etc. At the same time, our social activities are increasingly embedded in the ICT environments that enable and enhance our ability to transact, share experiences, and maintain social relationships. One of the best ways to explore these developments is through the mining and analysis of data, which are collected through mobile phones and allow us to investigate how individuals act when embedded in a technology-enabled environment. The MOTIf project builds on the analysis and modeling of geo-localized temporally detailed but fully anonymised mobile phone call networks. These datasets allow us to address the two scientific objectives about spatiotemporal patterns of service usage of anonymised individuals to learn when, where, and what people are doing; and about the fine-grained sociodemographic structure of society and its effect on the the individual social behaviour. In other words our goal in general is to understand how individuals behave in a dynamic techno-social environment.

7.4. International Research Visitors

7.4.1. Visits of International Scientists

Prof. Antonio F. Loureiro is a Visiting Researcher at Infine for 3 months, under the *DigiCosme Visiting Professor* funding. He worked with Aline C. Viana and the internship Joao Batista Borges on the inference of motifs from daily human mobility. He is also the Brazilian coordinator of the EMBRACE Inria associate team. He will give a series of lectures on *"What can a mobility trace tell us?"*.

7.4.1.1. Internships

Joao Batista Borges visited us for 2 weeks on October 2018 and will return on January 2019. The visits enter in the context of EMBRACE associated team. He work with Aline C. Viana and Antonio Loureiro on the extraction of motifs of mobility patterns of individuals that, when merged together, describe their daily motion and can be used to enhance mobility prediction.

Diego Madariaga Roman did an internship of 3 months at our team (Sep-Nov 2018). He work with Aline C. Viana, Marco Fiore and Panagiota Katsikouli on adaptive sampling frequency of human mobility.

Lucas Santos did an internship of 5 months at our team (May-Nov 2018), in the context of EMBRACE associated team. He work with Aline C. Viana and Pedro Olmo on the investigation of causalities in habits of human visits.

7.4.2. Visits to International Teams

7.4.2.1. Research Stays Abroad

Emmanuel Baccelli is Visiting Professor at Freie Universitaet (FU) Berlin, within the context of the formal collaboration IOTPUSH with this university on research topics about the Internet of Things, RIOT and Information-Centric Networking.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events Organisation

8.1.1.1. Member of the Organizing Committees

1. **Aline C. Viana** is Publicity co-Chair of: IEEE PICom 2018; IEEE MiSeNet 2018 (jointly with IEEE Infocom 2018); DCOSS 2018; DCOSS 2019; CoUrb 2019 (jointly with DCOSS 2019).
2. **Aline C. Viana** is Stutend Travel Grant co-chair of IEEE Infocom 2019.

8.1.2. Scientific Events Selection

8.1.2.1. Chair of Conference Program Committees

1. **Emmanuel Baccelli** was chair of the RIOT Summit 2018.

8.1.2.2. Member of the Conference Program Committees

1. **Aline C. Viana** is/was TPC member of: Altogel 2018; Algotel 2019; NTMS 2018; LADaS 2018.

8.1.3. Journal

8.1.3.1. Member of the Editorial Boards

1. **Aline C. Viana** is an Editorial Board member of Urban Computing Spring book series (http://www.springer.com/series/15552), since Feb. 2018; an Associate Editor of ACM SIGCOMM Computer Communication Review (ACM CCR), since May 2014; an Editorial Board member of Wireless Communications and Mobile Computing Open Access Journal of John Wiley&Sons and Hindawi since 2016.

8.1.3.2. Reviewer - Reviewing Activities

1. in 2018, **Emmanuel Baccelli** reviewed for ACM Transactions on IoT, ACM ICN, IEEE Transactions on Cloud Computing, IEEE Journal on Internet of Things, JSAC, PEMWN.
2. **Aline C. Viana** reviewed papers for ACM SIGCOMM CCR Journal, elsevier Pervasive and Mobile Computing Journal and Computer Communication Elsevier Journal, IEEE Trans. on Network and Service Management, on IEEE Internet Computing.

8.1.4. Standardization

1. **Emmanuel Baccelli and Cedric Adjih** have participated at several working groups at IETF during 2018.
2. **Emmanuel Baccelli and Cedric Adjih** have participated at several IETF hackathons during 2018.

8.1.5. Invited Talks

1. in 2018, **Emmanuel Baccelli** gave invited talks as EPO in Berlin, at Paris Open Source Summit, at the 2018 French-Japanese Workshop on Cybersecurity in Tokyo, and at OSIS 2018 in Paris.
2. in 2018, **Cedric Adjih** presented FIT IoT-LAB at the IoT Tunisia Forum 2018, described ICN IoT challenges at IoT Tunisia Workshop 2018, and in december, at a scientific seminar in University Paris 13.
3. **Aline C. Viana** was invited to give a talk at (1) LNCC, Petrópolis, Brazil, workshop on Urban Computing and Society in November 2018; (2) Univ. de la Rochelle, workshop CoFaBras 2018 in October 2018
4. **Aline C. Viana** was the keynote speaker of the CoUrb 2018 Workshop jointly organized with SBRC 2018, Brazil in May 2018. She will be a keynote at the GDR RSD Winter School in February 2019.

8.1.6. Research Administration

1. **Aline C. Viana** is the President of the Scientific Commission at Inria Saclay, responsible for the selection of candidates for the CORDI-S, Post-Doc and Delegation campaigns.
2. **Aline C. Viana** is the international coordinator of the EMBRACE associated team of Inria.
3. Together with the members of the team, **Aline C. Viana** has submitted the short proposal of the new team as well as has presented it at the BCEP and CEP. The long versio of the team proposal will be sent by next week.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

- Master : Emmanuel Baccelli, “IoT & Security”, summer semester seminar, Freie Universitaet Berlin.
- Master : Emmanuel Baccelli, “IoT & Security”, winter semester seminar, Freie Universitaet Berlin.
- Master : Emmanuel Baccelli, “Operating Systems for Small Connected Devices in the Internet of Things”, 3h cours magistral, Formation PESTO Corps des Mines, Telecom ParisTech, Paris France
- Engineering School: Cédric Adjih, “Internet of Thing”, 3h pratical courses, Telecom SudParis
- Engineering School: Cédric Adjih, “Internet of Thing”, 6h practical courses, ENSEA
- Lectures to under-graduation, masters and PhD students at foreign Universities: “Complete trajectory reconstruction from sparse mobile phone data”, 1h, Nov. 2018, LNCC, Petrópolis, Brazil
- Short-courses at conference: “Humans in the communication loop: enablers of next generation networks”, 4h, May 2017, SBRC 2018 conference, Brazil http://www.sbrc2018.ufscar.br/minicurso-1-mc-1/.

8.2.2. Supervision

- PHD Guangshuo Chen has defended his PhD in April 2018. *Thesis title*: Human Habits Investigation: from Mobility Reconstruction to Mobile Traffic Prediction. Ecole Polytechnique/ ED STIC Univ. Paris Saclay. Advisor: Aline C. Viana
- Roni Shigueta has defended his PhD in July 2018. *Thesis title:* Channel resource allocation of wireless interfaces of mobile devices.Co-tutelle: PUC-PR and Ecole Polytechnique/ ED STIC Univ. Paris Saclay. Advisors: Aline C. Viana and Mauro Fonseca (UTFPR, Brazil)
- PhD in progress: Licia Amichi, "Modelling explo- ration factor of human beings", since Sep. 2018.
- PhD in progress: Lucas Santos, "Investigating causalities in habits of human visits", since May 2018.
- PhD in progress: Douglas Teixeira, "Context-enhanced human predictability in short-term datasets with high spatial resolution", since April 2018.
- PhD in progress: Joao B. Borges, "Revealing motifs in human mobility", since Nov. 2018.
- PhD in progress: Rafael Costa, "Human-enhanced forward- ing strategies for Device-to-Device (D2D) communication", since Sep 2017.

8.2.3. Juries

- **Reviewer for PhD thesis committee:**Aline Viana was reviewer for the following PhDs: H. Chelle (INP/Univ. de Toulouse, Dec. 2018); Y. Zhou (CentraleSupélec, Jui.2018); F. Coriat (UPMC, Dec. 2018);
- **Examiner for PhD thesis committee:**Aline Viana was examiner for the committees of the following PhDs: A. Boubrima (INSA-Lyon/AGORA, Feb. 2019); L. Pajević (KTH, Nov. 2018);
- **Mid-term Examiner PhD thesis committee:**Emmanuel Baccelli was mid-term reviewer for the PhD of Benjamin Beurdouche (Inria, June 2018). Aline Viana was mid-term examiner for the following PhDs: M. Sardara (Cisco/TPT, Jun. 2018); M. Charfi and I. D. Adamou (CentraleSupélec, Jun. 2018); C. Bertier (Thales/UPMC, Avr. 2018);

9. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] G. CHEN.*Human Habits Investigation : from Mobility Reconstruction to Mobile Traffic Prediction*, Université Paris-Saclay, April 2018, https://pastel.archives-ouvertes.fr/tel-01784503

Articles in International Peer-Reviewed Journal

[2] I. AMDOUNI, C. ADJIH, N. AITSAADI, P. MUHLETHALER.*Extensive Experimentations on Opportunistic Routing in Wireless Sensor Networks*, in "Sensors", September 2018, vol. 18, n^o 9, 3031, https://hal.inria.fr/hal-01961054

[3] G. CHEN, S. HOTEIT, A. CARNEIRO VIANA, M. FIORE, C. SARRAUTE.*Enriching sparse mobility information in Call Detail Records*, in "Computer Communications", 2018, vol. 122, p. 44-58 [*DOI :* 10.1016/J.COMCOM.2018.03.012], https://hal.archives-ouvertes.fr/hal-01756120

Invited Conferences

[4] G. CHEN, A. CARNEIRO VIANA, M. FIORE.*Takeaways in Large-scale Human Mobility Data Mining*, in "IEEE International Symposium on Local and Metropolitan Area Networks", Washington, United States, June 2018, https://hal.inria.fr/hal-01795633

International Conferences with Proceedings

[5] E. BACCELLI, J. DOERR, O. JALLOULI, S. KIKUCHI, A. MORGENSTERN, F. A. PADILLA, K. SCHLEISER, I. THOMAS.*Reprogramming Low-end IoT Devices from the Cloud*, in "The 3rd Cloudification of the Internet of Things Conference (CIoT 2018)", Paris, France, July 2018, https://hal.inria.fr/hal-01960405

[6] E. BACCELLI, J. DOERR, S. KIKUCHI, F. A. PADILLA, K. SCHLEISER, I. THOMAS.*Scripting Over-The-Air: Towards Containers on Low-end Devices in the Internet of Things*, in "IEEE PerCom 2018", Athens, Greece, March 2018, p. 1-4, https://hal.inria.fr/hal-01766610

[7] L. DAUPHIN, E. BACCELLI, C. ADJIH.*RIOT-ROS2: Low-Cost Robots in IoT Controlled via Information-Centric Networking*, in "7th IFIP/IEEE International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks (PEMWN 2018)", Toulouse, France, September 2018, https://hal.inria.fr/hal-01898889

Conferences without Proceedings

[8] P. KATSIKOULI, A. CARNEIRO VIANA, M. FIORE, A. TARABLE.*L' étude de la fréquence d' échantillonnage des mouvements des humains*, in "CORES 2018 - Rencontres Francophones sur la Conception de Protocoles, l'Évaluation de Performance et l'Expérimentation des Réseaux de Communication", Roscoff, France, May 2018, p. 1-4, https://hal.archives-ouvertes.fr/hal-01784475

Other Publications

[9] B. ADAMSON, C. ADJIH, J. BILBAO, V. FIROIU, F. FITZEK, G. A. M. SAMAH, E. LOCHIN, A. MASUCCI, M.-J. MONTPETIT, M. V. PEDERSEN, G. PERALTA, V. ROCA, P. SAXENA, S. SIVAKUMAR.*Taxonomy of Coding Techniques for Efficient Network Communications*, June 2018, Internet Research Task Force, Request For Comments (RFC) 8406, https://datatracker.ietf.org/doc/rfc8406/, https://hal.inria.fr/hal-00998506

[10] V. ROCA, J. DETCHART, C. ADJIH, M. V. PEDERSEN.*Generic Application Programming Interface (API) for Sliding Window FEC Codes*, November 2018, p. 1-23, Internet Research Task Force - Working document of the Network Coding Research Group (NWCRG), draft-roca-nwcrg-generic-fec-api-04 (work in progress), https://datatracker.ietf.org/doc/draft-roca-nwcrg-generic-fec-api/, https://hal.inria.fr/hal-01630138

Project-Team LIFEWARE

Computational systems biology and optimization

RESEARCH CENTER
Saclay - Île-de-France

THEME
Computational Biology

Table of contents

Project-Team LIFEWARE

Creation of the Team: 2014 January 01, updated into Project-Team: 2015 April 01

Keywords:

Computer Science and Digital Science:

A2.1.1. - Semantics of programming languages
A2.1.5. - Constraint programming
A2.1.10. - Domain-specific languages
A2.2.1. - Static analysis
A2.3.2. - Cyber-physical systems
A2.4. - Formal method for verification, reliability, certification
A2.4.1. - Analysis
A2.4.2. - Model-checking
A2.4.3. - Proofs
A3.4.2. - Unsupervised learning
A3.4.4. - Optimization and learning
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.2. - Stochastic Modeling
A6.1.3. - Discrete Modeling (multi-agent, people centered)
A6.1.4. - Multiscale modeling
A6.2.4. - Statistical methods
A6.2.6. - Optimization
A6.3.1. - Inverse problems
A6.3.4. - Model reduction
A7.2. - Logic in Computer Science
A8.1. - Discrete mathematics, combinatorics
A8.2. - Optimization
A8.7. - Graph theory
A9.7. - AI algorithmics

Other Research Topics and Application Domains:

B1. - Life sciences
B1.1.2. - Molecular and cellular biology
B1.1.7. - Bioinformatics
B1.1.8. - Mathematical biology
B1.1.10. - Systems and synthetic biology
B2. - Health
B2.2.3. - Cancer
B2.4.1. - Pharmaco kinetics and dynamics
B2.4.2. - Drug resistance
B9. - Society and Knowledge

1. Team, Visitors, External Collaborators

Research Scientists

François Fages [Team leader, Inria, Senior Researcher, HDR]
Grégory Batt [Inria, Senior Researcher, HDR]
Jakob Ruess [Inria, Researcher]
Sylvain Soliman [Inria, Researcher, HDR]

Faculty Member

Philippe Dague [Univ Paris-Sud, Professor, until Sep 2018]

External Collaborator

Denis Thieffry [Ecole Normale Supérieure Paris, HDR]

Technical Staff

David Coudrin [Inria]
Steven Fletcher [Inria]
Matthieu Pichené [Inria, from Apr 2018 until Jul 2018]
Sebastian Ramon Sosa Carrillo [Inria, until Jan 2018]
François Bertaux [Institut Pasteur, from Apr 2018]
Eva Dechaux [Univ Paris-Saclay, until Aug 2018]
Andjela Davidovic [Institut Pasteur]

PhD Students

Chetan Aditya [Inria, from Feb 2018]
Virgile Andreani [Ecole polytechnique]
Eléonore Bellot [Inria, from Sep 2018]
Arthur Carcano [Univ. Paris-Diderot, from Oct 2018]
Julien Martinelli [INSERM, from Oct 2018]
Sebastian Ramon Sosa Carrillo [Inria, from Feb 2018]
Elise Weill Duflos [Inria, from Oct 2018]

Post-Doctoral Fellow

Mathieu Hemery [Inria, from Mar 2018]

Visiting Scientist

Lucia Nasti [Univ. of Pisa, Italy, from Nov 2018]

Administrative Assistants

Natalia Alves [Inria, from Jul 2018]
Adeline Lochet [Inria, from Jul 2018]
Corinne Petitot [Inria, until Jul 2018]

2. Overall Objectives

2.1. Overall Objectives

This project aims at developing formal methods and experimental settings for understanding the cell machinery and establishing computational paradigms in cell biology. It is based on the vision of **cells as machines, biochemical reaction networks as programs**, and on the use of concepts and tools from computer science to master the complexity of cell processes.

This project addresses fundamental research issues in computer science on the **interplay between structure and dynamics** in large interaction networks, and on **mixed analog-discrete computation**. We contribute to the theory of biochemical computation, and develop since 2002 a modelling, analysis and synthesis software, the Biochemical Abstract Machine, BIOCHAM. The reaction rule-based language of this system allows us to reason about biochemical reaction networks at different levels of abstraction, in the stochastic, differential, discrete, Boolean and hybrid semantics of reaction networks. We develop a variety of static analysis methods before going to simulations and dynamical analyses. We use **quantitative temporal logics** as a mean to formalise biological behaviours with imprecise data and to constrain model building or network synthesis.

A **tight integration between dry lab and wet lab efforts** is also essential for the success of the project. This is achieved through tight collaborations with biologists and experimentalists. Furthermore, half of Lifeware is in the InBio group at Institut Pasteur headed by Grégory Batt who develops an experimental platform for the closed-loop control of intracellular processes. This platform combines hardware (microfluidic device and microscope), software (cell tracking and model-based predictive control algorithms) and liveware (genetically modified living cells). The originality of this project thus also deals with the recourse to advanced microscopy and synthetic biology technologies to perform accurate observations, modifications and **real-time control** at both **single cell and cell population levels**.

For this to work, collaborations with top international leaders of these techniques have been established, and consolidated with student exchange programs, especially in the framework of the Doctorate School FIRE, Frontiers in Life Sciences, FdV to which we are affiliated, in addition to the Doctorate School Sciences et technologies de l'information et de la communication (STIC).

Because of the importance of optimization techniques in our research, we keep some activity purely dedicated to optimization problems, in particular on constraint programming methods for computing with partial information systems and solving NP-hard static analysis problems, and on continuous optimization methods for dealing with continuous parameters.

3. Research Program

3.1. Computational Systems Biology

Bridging the gap between the complexity of biological systems and our capacity to model and **quantitatively predict system behaviors** is a central challenge in systems biology. We believe that a deeper understanding of the concept and theory of biochemical computation is necessary to tackle that challenge. Progress in the theory is necessary for scaling, and enabling the application of static analysis, module identification and decomposition, model reductions, parameter search, and model inference methods to large biochemical reaction systems. A measure of success on this route will be the production of better computational modeling tools for elucidating the complex dynamics of natural biological processes, designing synthetic biological circuits and biosensors, developing novel therapy strategies, and optimizing patient-tailored therapeutics.

Progress on the **coupling of models to data** is also necessary. Our approach based on quantitative temporal logics provides a powerful framework for formalizing experimental observations and using them as formal specification in model building. Key to success is a tight integration between *in vivo* and *in silico* work, and on the mixing of dry and wet experiments, enabled by novel biotechnologies. In particular, the use of micro-fluidic devices makes it possible to measure behaviors at both single-cell and cell population levels *in vivo*, provided innovative modeling, analysis and control methods are deployed *in silico*.

In synthetic biology, while the construction of simple intracellular circuits has shown feasible, the design of larger, **multicellular systems** is a major open issue. In engineered tissues for example, the behavior results from the subtle interplay between intracellular processes (signal transduction, gene expression) and intercellular processes (contact inhibition, gradient of diffusible molecule), and the question is how should cells be genetically modified such that the desired behavior robustly emerges from cell interactions.

3.2. Chemical Reaction Network (CRN) Theory

Feinberg's chemical reaction network theory and Thomas's influence network analyses provide sufficient and/or necessary structural conditions for the existence of multiple steady states and oscillations in regulatory networks. Those conditions can be verified by static analyzers without knowing kinetic parameter values nor making any simulation. In this domain, most of our work consists in analyzing the interplay between the **structure** (Petri net properties, influence graph, subgraph epimorphisms) and the **dynamics** (Boolean, CTMC, ODE, time scale separations) of biochemical reaction systems. In particular, our study of influence graphs of reaction systems, our generalization of Thomas' conditions of multi-stationarity and Soulé's proof to reaction systems [0], the inference of reaction systems from ODEs [0], the computation of structural invariants by constraint programming techniques, and the analysis of model reductions by subgraph epimorphisms now provide solid ground for developing static analyzers, using them on a large scale in systems biology, and elucidating modules.

3.3. Logical Paradigm for Systems Biology

Our group was among the first ones in 2002 to apply **model-checking** methods to systems biology in order to reason on large molecular interaction networks, such as Kohn's map of the mammalian cell cycle (800 reactions over 500 molecules) [0]. The logical paradigm for systems biology that we have subsequently developed for quantitative models can be summarized by the following identifications :

biological model = transition system K

dynamical behavior specification = temporal logic formula ϕ

model validation = model-checking $K,\ s \models ?\ \phi$

model reduction = sub-model-checking, $K' \subset K$ s.t. $K'?,\ s \models \phi$

model prediction = formula enumeration, ϕ s.t. $K,\ s \models \phi?$

static experiment design = symbolic model-checking, state s s.t. $K,\ s? \models \phi$

model synthesis = constraint solving $K?,\ s \models \phi$

dynamic experiment design = constraint solving $K?,\ s? \models \phi$

In particular, the definition of a continuous satisfaction degree for **first-order temporal logic** formulae with constraints over the reals, was the key to generalize this approach to quantitative models, opening up the field of model-checking to model optimization [0] This line of research continues with the development of temporal logic patterns with efficient constraint solvers and their generalization to handle stochastic effects.

3.4. Computer-Aided Design of CRNs for Synthetic Biology

The continuous nature of many protein interactions leads us to consider models of analog computation, and in particular, the recent results in the theory of analog computability and complexity obtained by Amaury Pouly [0] and Olivier Bournez, establish fundamental links with digital computation. In a paper published last year [0] we have derived from these results the Turing completeness result of elementary CRNs (without polymerization) under the differential semantics, closing a long-standing open problem in CRN theory.

[0] Sylvain Soliman. A stronger necessary condition for the multistationarity of chemical reaction networks. Bulletin of Mathematical Biology, 75(11):2289–2303, 2013.

[0] François Fages, Steven Gay, Sylvain Soliman. Inferring reaction systems from ordinary differential equations. Journal of Theoretical Computer Science (TCS), Elsevier, 2015, 599, pp.64–78.

[0] N. Chabrier-Rivier, M. Chiaverini, V. Danos, F. Fages, V. Schächter. Modeling and querying biochemical interaction networks. Theoretical Computer Science, 325(1):25–44, 2004.

[0] On a continuous degree of satisfaction of temporal logic formulae with applications to systems biology A. Rizk, G. Batt, F. Fages, S. Soliman International Conference on Computational Methods in Systems Biology, 251-268

[0] Amaury Pouly, "Continuous models of computation: from computability to complexity", PhD Thesis, Ecole Polytechnique, Nov. 2015.

The proof of this result shows how computable function over the reals, described by Ordinary Differential Equations, namely by Polynomial Initial Value Problems (PIVP), can be compiled into elementary biochemical reactions, furthermore with a notion of analog computation complexity defined as the length of the trajectory to reach a given precision on the result. This opens a whole research avenue to analyze biochemical circuits in Systems Biology, transform behavioural specifications into biochemical reactions for Synthetic Biology, and compare artificial circuits with natural circuits acquired through evolution, from the novel point of view of analog computation and complexity.

3.5. Modeling of Phenotypic Heterogeneity in Cellular Processes

Since nearly two decades, a significant interest has grown for getting a quantitative understanding of the functioning of biological systems at the cellular level. Given their complexity, proposing a model accounting for the observed cell responses, or better, predicting novel behaviors, is now regarded as an essential step to validate a proposed mechanism in systems biology. Moreover, the constant improvement of stimulation and observation tools creates a strong push for the development of methods that provide predictions that are increasingly precise (single cell precision) and robust (complex stimulation profiles).

It is now fully apparent that cells do not respond identically to a same stimulation, even when they are all genetically-identical. This phenotypic heterogeneity plays a significant role in a number of problems ranging from cell resistance to anticancer drug treatments to stress adaptation and bet hedging.

Dedicated modeling frameworks, notably **stochastic** modeling frameworks, such as chemical master equations, and **statistic** modeling frameworks, such as ensemble models, are then needed to capture biological variability.

Appropriate mathematical and computational tools should then be employed for the analysis of these models and their calibration to experimental data. One can notably mention **global optimization** tools to search for appropriate parameters within large spaces, **moment closure** approaches to efficiently approximate stochastic models [0], and (stochastic approximations of) the **expectation maximization** algorithm for the identification of mixed-effects models [0].

3.6. External Control of Cell Processes

External control has been employed since many years to regulate culture growth and other physiological properties. Recently, taking inspiration from developments in synthetic biology, closed loop control has been applied to the regulation of intracellular processes. Such approaches offer unprecedented opportunities to investigate how a cell process dynamical information by maintaining it around specific operating points or driving it out of its standard operating conditions. They can also be used to complement and help the development of synthetic biology through the creation of hybrid systems resulting from the interconnection of in vivo and in silico computing devices.

In collaboration with Pascal Hersen (CNRS MSC lab), we developed a platform for gene expression control that enables to control protein concentrations in yeast cells. This platform integrates microfluidic devices enabling long-term observation and rapid change of the cells environment, microscopy for single cell measurements, and software for real-time signal quantification and model based control. We demonstrated in 2012 that this platform enables controlling the level of a fluorescent protein in cells with unprecedented accuracy and for many cell generations [0].

[0] Fages, François, Le Guludec, Guillaume and Bournez, Olivier, Pouly, Amaury. Strong Turing Completeness of Continuous Chemical Reaction Networks and Compilation of Mixed Analog-Digital Programs. In CMSB'17: Proceedings of the fiveteen international conference on Computational Methods in Systems Biology, pages 108–127, volume 10545 of Lecture Notes in Computer Science. Springer-Verlag, 2017.

[0] Moment-based inference predicts bimodality in transient gene expression, C. Zechner C, J. Ruess, P. Krenn, S. Pelet, M. Peter, J. Lygeros, and H. Koeppl, Proceedings of the National Academy of Sciences USA, 9(5):109(21):8340-5, 2012

[0] What population reveals about individual cell identity: estimation of single-cell models of gene expression in yeast, A. Llamosi, A.M. Gonzalez-Vargas, C. Versari, E. Cinquemani, G. Ferrari-Trecate, P. Hersen, and G. Batt, PLoS Computational Biology, 9(5): e1003056, 2015

More recently, motivated by an analogy with a benchmark control problem, the stabilization of an inverted pendulum, we investigated the possibility to balance a genetic toggle switch in the vicinity of its unstable equilibrium configuration. We searched for solutions to balance an individual cell and even an entire population of heterogeneous cells, each harboring a toggle switch [0].

Independently, in collaboration with colleagues from IST Austria, we investigated the problem of controlling cells, one at a time, by constructing an integrated optogenetic-enabled microscopy platform. It enables experiments that bridge individual and population behaviors. We demonstrated: (i) population structuring by independent closed-loop control of gene expression in many individual cells, (ii) cell–cell variation control during antibiotic perturbation, (iii) hybrid bio-digital circuits in single cells, and freely specifiable digital communication between individual bacteria [0].

3.7. Constraint Solving and Optimization

Constraint solving and optimization methods are important in our research. On the one hand, static analysis of biochemical reaction networks involves solving hard combinatorial optimization problems, for which **constraint programming** techniques have shown particularly successful, often beating dedicated algorithms and allowing to solve large instances from model repositories. On the other hand, parameter search and model calibration problems involve similarly solving hard continuous optimization problems, for which **evolutionary algorithms**, and especially the covariance matrix evolution strategy (CMA-ES) [0] have been shown to provide best results in our context, for up to 100 parameters. This has been instrumental in building challenging quantitative models, gaining model-based insights, revisiting admitted assumptions, and contributing to biological knowledge [00].

4. Application Domains

4.1. Preamble

Our collaborative work on biological applications is expected to serve as a basis for groundbreaking advances in cell functioning understanding, cell monitoring and control, and novel therapy design and optimization. Our collaborations with biologists are focused on **concrete biological questions**, and on the building of predictive models of biological systems to answer them. Furthermore, one important application of our research is the development of a **modeling software** for computational systems biology.

[0] Jannis Uhlendorf, Agnés Miermont, Thierry Delaveau, Gilles Charvin, François Fages, Samuel Bottani, Grégory Batt, Pascal Hersen. Long-term model predictive control of gene expression at the population and single-cell levels. Proceedings of the National Academy of Sciences USA, 109(35):14271–14276, 2012.

[0] Jean-Baptiste Lugagne, Sebastian Sosa Carrillo and Melanie Kirch, Agnes Köhler, Gregory Batt and Pascal Hersen. Balancing a genetic toggle switch by real-time feedback control and periodic forcing. Nature Communications, 8(1):1671, 2017.

[0] Remy Chait, Jakob Ruess, Tobias Bergmiller and Gavsper Tkavcik, Cvalin Guet. Shaping bacterial population behavior through computer-interfaced control of individual cells. Nature Communications, 8(1):1535, 2017.

[0] N. Hansen, A. Ostermeier (2001). Completely derandomized self-adaptation in evolution strategies. Evolutionary Computation, 9(2) pp. 159–195.

[0] Domitille Heitzler, Guillaume Durand, Nathalie Gallay, Aurélien Rizk, Seungkirl Ahn, Jihee Kim, Jonathan D. Violin, Laurence Dupuy, Christophe Gauthier, Vincent Piketty, Pascale Crépieux, Anne Poupon, Frédérique Clément, François Fages, Robert J. Lefkowitz, Eric Reiter. Competing G protein-coupled receptor kinases balance G protein and β-arrestin signaling. Molecular Systems Biology, 8(590), 2012.

[0] Pauline Traynard, Céline Feillet, Sylvain Soliman, Franck Delaunay, François Fages. Model-based Investigation of the Circadian Clock and Cell Cycle Coupling in Mouse Embryonic Fibroblasts: Prediction of RevErb-alpha Up-Regulation during Mitosis. Biosystems, 149:59–69, 2016.

4.2. Modeling software for systems biology and synthetic biology

Since 2002, we develop an open-source software environment for modeling and analyzing biochemical reaction systems. This software, called the Biochemical Abstract Machine (BIOCHAM), is compatible with SBML for importing and exporting models from repositories such as BioModels. It can perform a variety of static analyses, specify behaviors in Boolean or quantitative temporal logics, search parameter values satisfying temporal constraints, and make various simulations. While the primary reason of this development effort is to be able to **implement our ideas and experiment them quickly on a large scale**, BIOCHAM is used by other groups either for building models, for comparing techniques, or for teaching (see statistics in software section). BIOCHAM-WEB is a web application which makes it possible to use BIOCHAM without any installation. We plan to continue developing BIOCHAM for these different purposes and improve the software quality.

4.3. Coupled models of the cell cycle and the circadian clock

Recent advances in cancer chronotherapy techniques support the evidence that there exist important links between the cell cycle and the circadian clock genes. One purpose for modeling these links is to better understand how to efficiently target malignant cells depending on the phase of the day and patient characterictics. These questions are at the heart of our collaboration with Franck Delaunay (CNRS Nice) and Francis Lévi (Univ. Warwick, GB, formerly INSERM Hopital Paul Brousse, Villejuif) and of our participation in the ANR HYCLOCK project and in the submitted EU H2020 C2SyM proposal, following the former EU EraNet Sysbio C5SYS and FP6 TEMPO projects. In the past, we developed a coupled model of the Cell Cycle, Circadian Clock, DNA Repair System, Irinotecan Metabolism and Exposure Control under Temporal Logic Constraints [0]. We now focus on the bidirectional coupling between the cell cycle and the circadian clock and expect to gain fundamental insights on this complex coupling from computational modeling and single-cell experiments.

4.4. Biosensor design and implementation in non-living protocells

In collaboration with Franck Molina (CNRS, Sys2Diag, Montpellier) and Jie-Hong Jiang (NTU, Taiwan) we ambition to apply our techniques to the design and implementation of biosensors in non-living vesicles for medical applications. Our approach is based on purely protein computation and on our ability to compile controllers and programs in biochemical reactions. The realization will be prototyped using a microfluidic device at CNRS Sys2Diag which will allow us to precisely control the size of the vesicles and the concentrations of the injected proteins. It is worth noting that the choice of non-living chassis, in contrast to living cells in synthetic biology, is particularly appealing for security considerations and compliance to forthcoming EU regulation.

4.5. Functional characterization of the resistance of bacterial populations to antimicrobial treatments

Antibiotic resistance is becoming a problem of central importance at a global level. Two mechanisms are at the origin of non-susceptibility to antimicrobial treatments. The first one comes from adaptation of bacterial cells to antibacterial treatments, notably through the modification of efflux pumps or the expression of enzymes that degrade the antibiotics. Cells are individually resistant. The second one, typically found in resistances to β-lactams, a broad class of antibiotics, originates from the release in the environment of the antibiotic degrading enzymes by the dead cells. This leads to population effects by which cells become collectively resilient.

[0] Elisabetta De Maria, François Fages, Aurélien Rizk, Sylvain Soliman. Design, Optimization, and Predictions of a Coupled Model of the Cell Cycle, Circadian Clock, DNA Repair System, Irinotecan Metabolism and Exposure Control under Temporal Logic Constraints. Theoretical Computer Science, 412(21):2108 2127, 2011.

The functional characterization of these different effects is important for the best use of antibiotics (antibiotic stewardship). In collaboration with Lingchong You (Duke University) and with Philippe Glaser (Institut Pasteur), we develop experimental platforms, models, and optimal model calibration methods that gives precise estimations of individual resistance and collective resilience of bacterial populations to antibiotic treatments.

5. Highlights of the Year

5.1. Highlights of the Year

- **Multistationarity Analysis in CRNs**

  The non-existence of multiple steady states in very large chemical reaction networks, out of reach of symbolic computation methods, can be predicted by a remarkably fast graph rewriting algorithm, based on Soliman 2013's theorem [0]. Study published in the *Journal of Theoretical Biology* [1] (graphical abstract in Fig. 1).

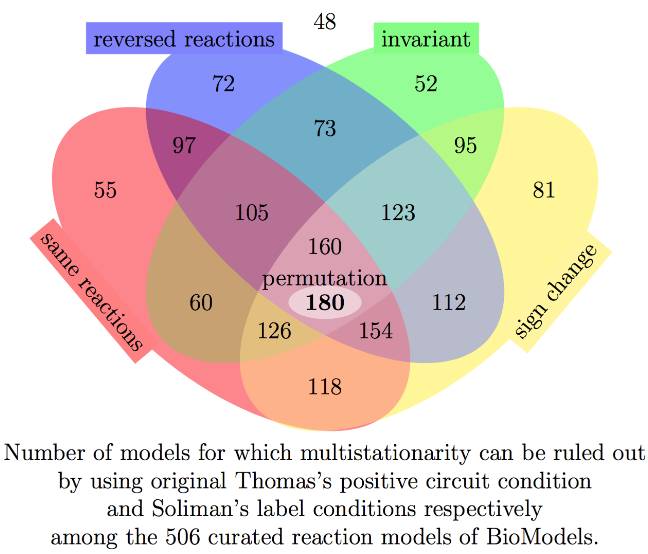


Figure 1. Graphical abstract of [1].

- **Distinguishing resistence from resilience to prolong antibiotic potency**

  Biomedical engineers at Duke University, in collaboration with Grégory Batt and Virgile Andréani, have shown experimentally that there is more than one flavor of antibiotic resistance and that it could – and should – be taken advantage of to keep first-line antibiotics in our medical arsenal. While an individual bacterium can be resistant to antibiotics, resilience only arises within a community. This happens when bacterial cells produce enough beta-lactamases to degrade the antibiotics, but not enough to save themselves from the initial onslaught. As some cells die and release more and more of the enzyme, the population as a whole eventually rids their environment of the antibiotic. Study published in *Science Advances* [6].

- **Biochemical Programs in Synthetic Cell-like Microreactors**

[0] Sylvain Soliman. A stronger necessary condition for the multistationarity of chemical reaction networks. Bulletin of Mathematical Biology, 75(11):2289–2303, 2013.

Researchers at Lab. CNRS-ALCEDIAG Sys2Diag in Montpellier, in collaboration with François Fages, have shown that an algorithm for the differential diagnosis of diabetes can be specified by three Boolean circuits and robustly implemented with real enzymes encapsulated in artificial vesicles that become fluorescent according to 5 different forms of diabetes. The robustness of the circuit was optimized in BIOCHAM by optimizing the initial concentrations of the enzymes with respect to a behavior specification in quantitative temporal logic. The protocells built with a microfluidic device were validated on a cohorte of patients' urines from Montpellier's Hospital. Study published in *Molecular Systems Biology* [3] (see Fig. 2).

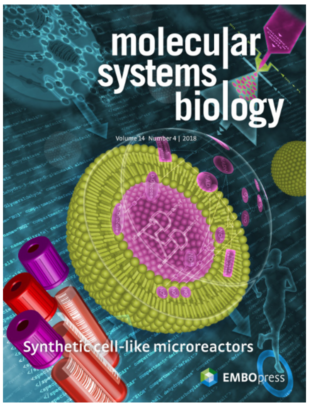


Figure 2. Artistic illustration by Courbet in cover page of Molecular Systems Biology [3].

5.1.1. Awards

- **La Recherche magazine 2019 Award - mention Information Sciences**

  The article[6] "Strong Turing Completeness of Continuous Chemical Reaction Networks and Compilation of Mixed Analog-Digital Programs" by F. Fages, G. Le Guludec, O. Bournez and A. Pouly, presented and awarded Best Paper at CMSB'17 last year has received the 2019 Award of magazine "La Recherche" - in Information Sciences.

6. New Software and Platforms

6.1. BIOCHAM

The Biochemical Abstract Machine

KEYWORDS: Systems Biology - Bioinformatics

FUNCTIONAL DESCRIPTION: The Biochemical Abstract Machine (BIOCHAM) is a software environment for modeling, analyzing and synthesizing biochemical reaction networks (CRNs) with respect to a formal specification of the observed or desired behavior of a biochemical system. BIOCHAM is compatible with the Systems Biology Markup Language (SBML) and contains some unique features about formal specifications in quantitative temporal logic, sensitivity and robustness analyses and parameter search in high dimension w.r.t. behavioral specifications, static analyses, and synthesis of CRNs.

RELEASE FUNCTIONAL DESCRIPTION: influence networks with forces – PAC learning of influence networks from time series data – synthesis of continuous reaction networks for mathematical functions defined by polynomial differential equations – complete modular rewriting of Biocham in SWI-Prolog

- Participants: François Fages, David Coudrin, Sylvain Soliman and Thierry Martinez
- Contact: François Fages
- URL: http://lifeware.inria.fr/biocham4/

6.2. Platforms

6.2.1. Smart experimental platforms to automate microbiology experiments

Models play a central role in our work, either to test our understanding or to guide the design of novel systems. Model development and parameter calibration necessitate informative experiments. We develop methods to assist with the optimal design of experiments. In consequence, we have to perform, in sequence or in parallel, experiments with possibly complex input profiles. This led us to develop experimental platforms that allow for flexible and automated stimulations and measurements. Three platforms are being developed, based on (i) a microplate photometer, (ii) a bioreactor platform coupled with a flow cytometer, and (iii) a microscope equipped with microfluidic systems, respectively. In all cases, the real-time measurement and actuation capabilities allow for making reactive experiments, notably including real-time control experiments.

7. New Results

7.1. Graphical Requirements for Multistationarity in CRNs and their Verification in BioModels

Participants: Adrien Baudier, François Fages, Sylvain Soliman.

Thomas's necessary conditions for the existence of multiple steady states in gene networks have been proved by Soulé with high generality for dynamical systems defined by differential equations. When applied to (protein) reaction networks however, those conditions do not provide information since they are trivially satisfied as soon as there is a bimolecular or a reversible reaction. Refined graphical requirements have been proposed to deal with such cases. In [1], we present for the first time a graph rewriting algorithm for checking the refined conditions given by Soliman, and evaluate its practical performance by applying it systematically to the curated branch of the BioModels repository. This algorithm analyzes all reaction networks (of size up to 430 species) in less than 0.05 second per network, and permits to conclude to the absence of multistationarity in 160 networks over 506. The short computation times obtained in this graphical approach are in sharp contrast to the Jacobian-based symbolic computation approach. We also discuss the case of one extra graphical condition by arc rewiring that allows us to conclude on 20 more networks of this benchmark but with a high computational cost. Finally, we study with some details the case of phosphorylation cycles and MAPK signalling models which show the importance of modelling the intermediate complexations with the enzymes in order to correctly analyze the multistationarity capabilities of such biochemical reaction networks.

7.2. Influence Networks compared with CRNs: Semantics, Expressivity and Attractors

Participants: François Fages, Thierry Martinez [former member], David Rosenblueth [former member], Sylvain Soliman, Denis Thieffry.

Biochemical reaction networks are one of the most widely used formalism in systems biology to describe the molecular mechanisms of high-level cell processes. However modellers also reason with influence diagrams to represent the positive and negative influences between molecular species and may find an influence network useful in the process of building a reaction network. In [4], we introduce a formalism of influence networks with forces, and equip it with a hierarchy of Boolean, Petri net, stochastic and differential semantics, similarly to reaction networks with rates. We show that the expressive power of influence networks is the same as that of reaction networks under the differential semantics, but weaker under the discrete semantics. Furthermore, the hierarchy of semantics leads us to consider a (positive) Boolean semantics without test for absence, that we compare with the (negative) Boolean semantics with test for absence of gene regulatory networks à la Thomas. We study the monotonicity properties of the positive semantics and derive from them an algorithm to compute attractors in both the positive and negative Boolean semantics. We illustrate our results on models of the literature about the p53/Mdm2 DNA damage repair system, the circadian clock, and the influence of MAPK signaling on cell-fate decision in urinary bladder cancer.

As an application, in [11] methods are shown to add dynamics to large molecular influence maps.

7.3. Reducing CRNs by Tropicalization

Participants: Eléonore Bellot, François Fages, Aymeric Quesne, Sylvain Soliman, Eliott Suits.

We have shown in the past that model reduction relationships between CRNs can be detected on a large scale by the graph matching notion of subgraph epimorphism [0], furthermore quite efficiently with constraint programming or SAT solving techniques. However this approach does not allow us to actually reduce models. In the framework of the ANR-DFG SYMBIONT project [10] we are investigating model reduction methods based on tropicalization and constraint programming techniques [0] together with correctness conditions based on Tikhonov theorem.

7.4. Compiling mathematical functions and programs in CRNs

Participants: Auriane Cozic, Elisabeth Degrand, François Fages, Mathieu Hemery, Wei-Chih Huang, Lena Le Quellec, Sylvain Soliman.

In a previous paper, we have proven that any computable function over the reals in the sense of computable analysis (i.e. computable with finite yet arbitrary precision by a Turing machine) is computable by a continuous CRN over a finite set of molecular species. In this approach, the real-valued molecular concentrations are the information carriers and computation can be purely analog. We have derived from the proof of this result a compiler of real functions (of either time or input concentrations) specified by polynomial initial value problems (PIVP) in elementary CRNs. This compiler makes it possible to automate the design of abstract CRNs for implementing arbitrary computable functions over the reals presented by PIVPs, in particular arithmetic, trigonometric, sigmoid and logical functions. The compilation of sequentiality, program control flows and mixed analog-digital imperative programs lead us however to consider more efficient implementations of Heavyside functions with simple CRNs that have no simple mathematical expression as input/output functions. Our goal is to develop a compiler of high-level mixed analog-digital programs in efficient abstract CRNs amenable to practical implementation with real enzymes in DNA-free vesicles, as illustrated in Section 7.8.

7.5. Evolving CRNs from data time series

Participants: Elisabeth Degrand, François Fages, Jérémy Grignard [former&future Member], Mathieu Hemery, Sylvain Soliman.

Another approach to CRN design is by evolutionary algorithms. Given a function given by its graph with a finite set of points, and using the same framework based on PIVPs as above, we have designed a genetic algorithm which interleaves the evolution of a population of PIVPs with parameter optimization using CMA-ES for fitting the input curve. On the cosine function, this algorithm recovers PIVPs equivalent to the standard PIVP for cosine, while on Heavyside functions, the algorithm finds (mathematically mysterious) CRNs that are much simpler than Hill functions of high order for instance.

[0] Steven Gay, François Fages, Thierry Martinez, Sylvain Soliman, Christine Solnon. On the subgraph Epimorphism Problem. Discrete Applied Mathematics, 162:214–228, 2014.

[0] Sylvain Soliman, François Fages, Ovidiu Radulescu. A constraint solving approach to model reduction by tropical equilibration. Algorithms for Molecular Biology, 9(24), 2014.

7.6. Learning CRNs from data time series

Participants: François Fages, Jérémy Grignard [former&future Member], Nicolas Levy, Julien Martinelli, Sylvain Soliman.

The problem of learning a mechanistic model from observed data is more difficult than learning a blackbox model fitting the data, due to the difference between causal relationships and correlations. In a biological context, learning a mechanistic model from experimental data would help understanding the underlying biological processes. To that end, considering multiple time series data generated by a hidden CRN from different initial states (by either stochastic or differential simulation), we develop a clustering-based algorithm for the inference of biological reaction networks. The output is a set of reactions which can be used to generate new traces. A model selection method is derived from these newly generated traces. We evaluate the performance of this algorithm on a range of models from Biomodels.

7.7. Optimizing CRN robustness

Participants: François Fages, Lucia Nasti, Sylvain Soliman.

In [7] we present two complementary notions of robustness of a system with respect to a property of its behaviour expressed in temporal logic: first the statistical notion of model robustness to parameter perturbations, defined as its mean functionality; and second, a metric notion of formula satisfaction robustness, defined as the penetration depth in the validity domain of the temporal logic constraints. We show how the formula robustness can be used in BIOCHAM-4 with no extra cost as an objective function in the parameter optimization procedure, to actually improve CRN robustness. We illustrate these unique features with a classical example of the hybrid systems community and provide some performance figures on a model of MAPK signalling with 37 parameters.

7.8. Robust biochemical programming of synthetic microreactors

Participants: Auriane Cozic, François Fages, Wei-Chih Huang, Lena Le Quellec, Lucia Nasti, Sylvain Soliman.

Biological systems have evolved efficient sensing and decision-making mechanisms to maximize fitness in changing molecular environments. Synthetic biologists have exploited these capabilities to engineer control on information and energy processing in living cells. While engineered organisms pose important technological and ethical challenges, de novo assembly of non-living biomolecular devices could offer promising avenues towards various real-world applications. However, assembling biochemical parts into functional information processing systems has remained challenging due to extensive multidimensional parameter spaces that must be sampled comprehensively in order to identify robust, specification compliant molecular implementations. In [3], we introduce a systematic methodology based on automated computational design and microfluidics enabling the programming of synthetic cell-like microreactors embedding biochemical logic circuits, or protosensors, to perform accurate biosensing and biocomputing operations in vitro according to temporal logic specifications. We show that proof-of-concept protosensors integrating diagnostic algorithms detect specific patterns of biomarkers in human clinical samples. Protosensors may enable novel approaches to medicine and represent a step towards autonomous micromachines capable of precise interfacing of human physiology or other complex biological environments, ecosystems or industrial bioprocesses.

7.9. Identification of individual cells from z-stacks of bright-field microscopy images

Participants: Grégory Batt, Chiara Fracassi [former Member], Jean-Baptiste Lugagne [former Member].

Obtaining single cell data from time-lapse microscopy images is critical for quantitative biology, but bottlenecks in cell identification and segmentation must be overcome. In [5], we propose a novel, versatile method that uses machine learning classifiers to identify cell morphologies from z-stack bright-field microscopy images. We show that axial information is enough to successfully classify the pixels of an image, without the need to consider in focus morphological features. This fast, robust method can be used to identify different cell morphologies, including the features of *E. coli, S. cerevisiae* and epithelial cells, even in mixed cultures. Our method demonstrates the potential of acquiring and processing Z-stacks for single-layer, single-cell imaging and segmentation.

7.10. Applying ecological resistance and resilience to dissect bacterial antibiotic responses

Participants: Virgile Andreani, Grégory Batt.

An essential property of microbial communities is the ability to survive a disturbance. Survival can be achieved through resistance, the ability to absorb effects of a disturbance without a notable change, or resilience, the ability to recover after being perturbed by a disturbance. These concepts have long been applied to the analysis of ecological systems, although their interpretations are often subject to debate. In [6], we show that this framework readily lends itself to the dissection of the bacterial response to antibiotic treatment, where both terms can be unambiguously defined. The ability to tolerate the antibiotic treatment in the short term corresponds to resistance, which primarily depends on traits associated with individual cells. In contrast, the ability to recover after being perturbed by an antibiotic corresponds to resilience, which primarily depends on traits associated with the population. This framework effectively reveals the phenotypic signatures of bacterial pathogens expressing extended-spectrum β-lactamases when treated by a β-lactamase antibiotic. Our analysis has implications for optimizing treatment of these pathogens using a combination of a β-lactamase and a β-lactamase inhibitor. In particular, our results underscore the need to dynamically optimize combination treatments based on the quantitative features of the bacterial response to the antibiotic or the Bla inhibitor.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR Projects

- ANR-FWF CyberCircuits (2018-2022): "Cybergenetic circuits to test composability of gene networks", co-coordinated by C. Guet (IST Austria, Klosterneuburg, Austria) and J. Ruess (Inria EPI Lifeware);
- ANR-DFG SYMBIONT (2018-2021) on "Symbolic Methods for Biological Systems", coordinated by T. Sturm (CNRS, LORIA, Nancy, France) and A. Weber (Univ. Bonn, Germany) with F. Fages and F. Boulier (U. Lille), O. Radulescu (U. Montpellier), A. Schuppert (RWTH Aachen), S. Walcher (RWTH Aachen), W. Seiler (U. Kassel);
- ANR-MOST BIOPSY (2016-2020) on "Biochemical Programming System", coordinated by F. Molina (CNRS, Sys2diag, Montpellier) and J.H. Jiang (National Taiwan University), with F. Fages;
- ANR MEMIP (2016-2020) on "Mixed-Effects Models of Intracellular Processes", coordinated by G. Batt, with P. Hersen, (CNRS/Paris7), E. Cinquemani (Inria EPI IBIS) and M. Lavielle (Inria/CNRS/Polytechnique, EPI XPOP);
- ANR COGEX (2016-2019) on "Computer Aided Control of Gene Expression" coordinated by P. Hersen (MSC lab, CNRS/Paris7), with G. Batt and G. Truan (LISBP, CNRS/INSA);
- ANR Blanc HYCLOCK (2014-2018) on "Hybrid modeling of time for Circadian Clock Biology and Chronopharmacology", coordinated by F. Delaunay (CNRS, Nice), with F. Lévi (INSERM Paris-Sud), G. Bernot (CNRS I3S, Nice), O. Roux (Ecole Centrale Nantes), F. Fages and S. Soliman;
- ANR Blanc STOCH-MC (2014-2018) on "Stochastic Models: Scalable Model Checking", coordinated by Blaise Genest (Inria Rennes), with Grégory Batt, Wieslaw Zielonka (LIAFA), and Hugo Gimbert (LaBRI).

8.1.2. Inria Project Lab

- IPL COSY (2017-2021) “real-time control of synthetic microbial communities”, coordinated by Eugenio Cinquemani (Ibis, Inria), with Jean-Luc Gouzé (Biocore, Inria), Gregory Batt, Frédéric Bonnans (Commands, Inria), Efimov Denis (Non-A, Inria), and Hans Geiselmann (BIOP, Université Grenoble-Alpes), Beatrice Laroche (Maiage, Inra Jouy-en-Josas).

8.2. European Initiatives

8.2.1. H2020 Projects

- H2020 FET-OPEN COSY-BIO (2017-2020), “Control Engineering of Biological Systems for Reliable Synthetic Biology Applications”, coordinated by Diego di Bernardo (Tigem), with Filippo Menolascina (Edinburgh U), Mario di Bernardo (Naples U), Pascal Hersen (Paris7 U), Mustafa Khammash (ETHZ), Gregory Batt, Guy-Bart Stan (Imperial College), and Lucia Marucci (Bristol U).

8.3. International Research Visitors

8.3.1. Visits of International Scientists

The following researchers have been invited for short visits

- Carlo Spaccasassi, Microsoft Research Cambridge, UK
- Debdas Paul, Univ. Stuttgart, Germany

8.3.1.1. Internships

Lucia Nasti, PhD candidate at the Universita of Pisa, Italy, is visiting our group for 4 months.

8.3.2. Visits to International Teams

8.3.2.1. Research Stays Abroad

Jakob Ruess stayed at IST Austria twice a week in Feb and Nov 2018.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. Member of the Organizing Committees

Philippe Dague and François Fages were co-organizers of the Workshop on Computational Systems Biology for Cancer CSBC, Institut des Systèmes Complexes, Paris, France, 24-26 jan. 2018.

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

- Gregory Batt was member of the scientific program committee of
  - the 19th International Conference on Systems Biology (ICSB 2018), Oct 28 - Nov 1, Lyon, France
- François Fages was member of the program committee of
  - VEMDP'18 Verification of Engineered Molecular Devices and Programs. July 19, 2018, Oxford, United Kingdom. An affiliated workshop of CAV 2018.
  - CMSB'18 16th International Conference on Computational Methods in Systems Biology. 12th-14th September 2018, Faculty of Informatics, Masaryk University, Brno, CZ.
  - MCU'18 8th International Conference on Machines, Computations and Universality, June 28-30, 2018, Fontainebleau, France.
  - IJCAR'18 9th International Joint Conference on Automated Reasoning, part of the Federated Logic Conference FLOC'18, July 14-17, 2018, Oxford, UK.
  - DataMod'18 7th International Symposium "From Data to Models and Back", 25-26 June 2018, Toulouse, France
  - WCB'18 13th International Workshop on Constraint-based Methods for Bioinformatics, Oxford co-located with FLOC'18, July 2018.

9.1.2.2. Reviewer

- Jakob Ruess has reviewed scientific articles for
  - the 57th IEEE Conference on Decision and Control (CDC 2018), Dec. 17-19, Miami, USA;
  - the 17th European Control Conference (ECC 2019), June 25-28, Naples, Italy;
  - the Bioinformatics conference, Feb 22-24 2019, Prague, Czech Republic.
- Sylvain Soliman reviewed papers for
  - IJCAR and CMSB.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- François Fages is member of
  - the Editorial Board of the Computer Science area of the Royal Society Open Science journal, since 2014;
  - the Editorial Board of the journal RAIRO OR Operations Research, since 2004.

9.1.3.2. Reviewer - Reviewing Activities

In addition to their Editorial Board and Program Committee duties,

- Grégory Batt reviewed journal articles for *ACS Synthetic Biology, Cell Systems, Nature Communications, and Current Opinion in Systems Biology*;
- François Fages reviewed journal articles for *PLOS Computational Biology, Interface Focus, Natural Computing, BMC Systems Biology* and the book *Automated Reasoning for Systems Biology and Medicine*;
- Jakob Ruess reviewed journal articles for *PLOS Computational Biology, Journal of Mathematical Biology, Journal of Computational Physics, Processes, and Entropy*;
- Sylvain Soliman reviewed journal articles for *Journal of Theoretical Biology* and *Transactions in Computational Biology and Bioinformatics*.

9.1.4. Invited Talks

- Virgile Andreani gave an invited talk on *Modèles de résistance bactérienne aux antibiotiques to Journée apprentissage de modèles statistiques et stochastiques à partir de données biologiques*, Mar 2018, Rennes, France;
- Grégory Batt gave invited talks on
  - *Balancing a genetic toggle switch by real-time control and periodic stimulations*, IEEE Conference on Control Technology and Applications (CCTA 2018), Aug 21-24 2018, Copenhagen, Denmark
  - *Impact of phenotypic variability on commitment to apoptosis, Computational Systems Biology of Cancer*, Sept 2018, Paris
  - *Balancing a genetic toggle switch by real-time control and periodic stimulations*, iSSB / Genoscope seminar, Dec 2018, Evry, France
- François Fages gave invited talks on
  - *Vers une informatique de la cellule: programmes biochimiques naturels et synthétiques*, Conférence de rentrée de l'ENS Paris-Saclay (3h), Sep. 2018;
  - *Modeling and Design of Biological Systems*, Discrete Models and Formal Verification in Biology, Cambridge, UK, août 2018;
  - *Modeling and Design of Biological Systems*, Summer school on New strategies in medical diagnosis and personalized medicine, Ninh Binh, Vietnam, Jul. 2018;
  - *Computer-Aided Biochemical Programming of Synthetic Micro-reactors as Diagnostic Devices*, Groupe de Travail sur la biologie systémique symbolique, GT BIOSS, Marseille, Jul. 2018;
  - *Turing Completeness of Continuous Chemical Reaction Networks, an Informatic Perspective to Systems Biology and Synthetic Biology*, MODELIFE days, Université Provence Alpes Côte d'Azur, Nice, June 2018;
  - *A Constraint-based Implementation of Tropical Equilibrations in BIOCHAM towards integrating Stability Conditions for Computing Correct Model Reductions*, Scientific Meeting, Bonn, Allemagne, March 2018;
  - *From Turing Completeness of Continuous CRN to Abstract CRN Synthesis Workflow*, Formal Methods for the Synthesis of Biomolecular Circuits, Dagstuhl, Allemagne, Feb. 2018;
  - *Coupled Models of the Cell Cycle and Circadian Clock for Chronotherapy Optimization*, Computational Systems Biology for Cancer, Paris, Jan. 2018.

  Jakob Ruess gave invited talks on
  - *Virtual reality for bacteria*, C3BI Seminar, Feb 2018, Paris, France.
  - *Shaping bacterial population behavior through computer-interfaced control of individual cells*, Symposium on "Cybergenetics-at the interface between living and non-living regulatory systems" at the Annual Meeting of the German Association for General and Applied Microbiology (VAAM), Apr 2018, Wolfsburg, Germany.

9.1.5. Leadership within the Scientific Community

- Grégory Batt is
  - co-responsible of the working group on Symbolic Systems Biology (GT Bioss), gathering >150 researchers in 25 research teams; and member of
  - the Technical Committee on Systems Biology of IEEE and CSS societies;
  - the scientific board of the French research network on Systems and Synthetic Biology (GdR BioSynSys), gathering 40 labs and 300 researchers;

- the scientific board of the French research network on Bioinformatics (GdR BIM), gathering 56 labs and several hundreds of researchers;
- the scientific committee of the Advanced Course on Computational Systems Biology summer school, Aussois, 2019;
- the scientific committee of the Curie International Course on Systems Biology of Cancer, Paris, 2018.

- François Fages is member of
  - the Steering Committee of the International Conference on Computational Methods in Systems Biology, CMSB, since 2008.
  - the Think Tank of the structuring program MODELIFE "Modélisation physique et Mathématique du vivant", Université Provence Alpes Côte d'Azur.

9.1.6. Scientific Expertise

- Grégory Batt
  - was responsible for the Predictive Systems Biology chapter of the 2018-2022 strategic plan of Inria;
  - was reviewer and panel member of the ITMO Cancer/Inserm call on single-cell approaches to cancer;
  - and reviewer for the MIT International Science & Technology Initiatives (MISTI) Global Seed Funds and for the Leverhulme Trust.
- François Fages
  - was member of the Comité de sélection Maître de Conférence, ENS Paris-Saclay;
  - of the thesis advisory committee of one PhD student;
  - and evaluator of one European Research Council ERC Consolidator Grant;
  - one European Research Council ERC Starting Grant;
  - and one DigiCosme, Univ. Paris-Saclay project.
- Jakob Ruess has been a member of thesis advisory committee for two PhD students.

9.1.7. Research Administration

- François Fages is member of the "Comité des Projets du centre" Inria Saclay-IdF
- Sylvain Soliman is member of the "Commission Scientifique" of Inria Saclay-IdF

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master: Grégory Batt (coordinator and teacher: 35h) and Jakob Ruess (25h), *Computational Biology*, M1, Master Approches Interdisciplinaires du Vivant (AIV).

Master: Grégory Batt (3h) *Synthetic Biology and Control course in Molecular and Cellular Biology* Sorbonne Université, Paris.

Master: François Fages (coordinator module 24h and teacher 12h) *C2-19 Biochemical Programming*, Master Parisien de Recherche en Informatique (MPRI), Paris.

Master: François Fages (co-coordinator module 36h and teacher 18h) and Sylvain Soliman (co-coordinator, teacher 18h) INF555 - Constraint-based Modeling and Algorithms for Decision Making ProblemsMaster Artificial Intelligence, Ecole Polytechnique.

Master Modélisation en Pharmacologie Clinique et Epidémiologie Villejuif. François Fages (3h) Méthodes de modélisation informatique des processus cellulaires: cycle cellulaire et horloge circadienne, Bachelor 2: Eléonore Bellot (teacher 64h) *CSE201 Object-oriented Programming in C++* TD and project supervision

Thematic Research School: Advances in systems and synthetic biology, Evry. François Fages and Sylvain Soliman (3h) *BIOCHAM Workshop*

9.2.2. Supervision

PhD in progress: Virgile Andreani, Calibration efficace de modèles de résistance bactérienne aux antibiotiques à l'aide d'un plan d'expériences optimal, ED STIC, Ecole Polytechnique, Sept. 2016, Gregory Batt

PhD in progress: Sebastian Sosa, "Understanding cost of protein production in yeast", ED FdV, Université Sorbonne Paris Cité, Feb. 2018, Gregory Batt

PhD in progress: Chetan Aditya, "Control of heterogenous synthetic microbial systems", ED FdV, Université Sorbonne Paris Cité, Feb. 2018, Gregory Batt

PhD in progress: Elise Weill, "Optimal control of partial differential equation systems: Application to heterogeneous cell populations", ED STIC, Ecole Polytechnique, Oct. 2018, Frédéric Bonnans and Gregory Batt

PhD in progress: Arthur Carcano, "Iterative design of single-cell experiments to learn single-cell models of biological systems", ED FdV, Université Sorbonne Paris Cité, Oct. 2018, Jakob Ruess and Gregory Batt

PhD in progress : Eléonore Bellot, "Réduction de modèles différentiels par résolution de contraintes d'algèbre tropicale (min,+)", ED STIC, Ecole Polytechnique, Sept. 2018, F. Fages & S. Soliman (50-50%)

PhD in progress : Jérémy Grignard, "Apprentissage de modèles à partir de données pour la conception d'expériences de criblage et la recherche de médicaments", ED STIC, Ecole Polytechnique, dec. 2018, F. Fages & T. Dorval, Servier (50-50%)

PhD in progress : Julien Martinelli, "Apprentissage de modèles mécanistes à partir de données temporelles, application à la personnalisation de la chronothérapie des cancers", ED STIC, Ecole Polytechnique, oct. 2018, F. Fages & A. Ballesta, Inserm (50-50%)

Master's Thesis in progress : Elisabeth Degrand, "Evolving Chemical Reaction Networks", KTH Stockholm, Sweden, F. Fages & M. Hemery (50-50%)

9.2.3. Juries

- Grégory Batt participated in the jurys of
  - HdR of Benjamin Pfeuty, Université de Lille, *Rapporteur*, 23/11/2018
  - PhD of Arnaud Bonnaffoux, ENS de Lyon, *Président du Jury*, 12/10/2018
  - PhD of Alexandre Deloupy, Sorbonne Université, *Rapporteur*, 14/12/2018
- François Fages participated in the jurys of
  - HDR Thi Bich Han Dao, Université d'Orléans, *Président du jury*, 14 mars 2018.
  - PhD Alexandre Rocca, Université Grenoble, *Rapporteur*, 7 mai 2018.
  - PhD Benjamin Miraglio, Université de Nice, *Rapporteur*, 16 février 2018.
- Jakob Ruess was member of the jury of the PhD thesis of Mathieu Pichené, 25/06/2018.

9.3. Popularization

9.3.1. Articles and contents

Our publication on distinguishing resistance from resilience to antibiotic treatments [6] has been the object of a press release, institutional communications (at Inria and at Institut Pasteur), and some press attention, including notably by the UK daily mail journal.

9.3.2. Interventions

- Eléonore Bellot and Elise Weill participated at Fête de la Science in October 2018 at Inria Saclay île-de-France.
- Eléonore Bellot has received college school students for a visit to our research team on the design of logical gates in synthetic biology and their simulation in BIOCHAM.
- François Fages has received college school students for a visit to our research team with the question “Can we program any function ?”, answered negatively by proving Cantor’s theorem.

9.3.3. Creation of media or tools for science outreach

- François Fages has created several BIOCHAM notebooks associated to his course at MPRI and available online.
- François Fages and Sylvain Soliman have created one BIOCHAM tutorial notebook presented at ASSB 2018 and available online.

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] A. BAUDIER, F. FAGES, S. SOLIMAN.*Graphical Requirements for Multistationarity in Reaction Networks and their Verification in BioModels*, in "Journal of Theoretical Biology", December 2018, vol. 459, p. 79–89, https://arxiv.org/abs/1809.08891 [*DOI :* 10.1016/J.JTBI.2018.09.024], https://hal.archives-ouvertes.fr/hal-01879735

[2] F. BOULIER, F. FAGES, O. RADULESCU, S. S. SAMAL, A. SCHUPPERT, W. M. SEILER, T. STURM, S. WALCHER, A. WEBER.*The SYMBIONT project: symbolic methods for biological networks*, in "ACM Communications in Computer Algebra", September 2018, vol. 52, n^o 3, p. 67-70 [*DOI :* 10.1145/3313880.3313885], https://hal.inria.fr/hal-02061018

[3] A. COURBET, P. AMAR, F. FAGES, E. RENARD, F. MOLINA.*Computer-aided biochemical programming of synthetic microreactors as diagnostic devices*, in "Molecular Systems Biology", April 2018, vol. 14, n^o 4 [*DOI :* 10.15252/MSB.20177845], https://hal.inria.fr/hal-01779791

[4] F. FAGES, T. MARTINEZ, D. A. ROSENBLUETH, S. SOLIMAN.*Influence Networks compared with Reaction Networks: Semantics, Expressivity and Attractors*, in "IEEE/ACM Transactions on Computational Biology and Bioinformatics", 2018, vol. PP, n^o 99, p. 1-14 [*DOI :* 10.1109/TCBB.2018.2805686], https://hal.inria.fr/hal-01510216

[5] J.-B. LUGAGNE, S. JAIN, P. IVANOVITCH, Z. BEN MERIEM, C. VULIN, C. FRACASSI, G. BATT, P. HERSEN.*Identification of individual cells from z-stacks of bright-field microscopy images*, in "Scientific Reports", July 2018, vol. 8, n^o 1, 11455 [*DOI :* 10.1038/S41598-018-29647-5], https://hal.archives-ouvertes.fr/hal-01898065

[6] H. MEREDITH, V. ANDREANI, H. MA, A. LOPATKIN, A. LEE, D. ANDERSON, G. BATT, L. YOU.*Applying ecological resistance and resilience to dissect bacterial antibiotic responses*, in "Science Advances ", December 2018, vol. 4, n^o 12, eaau1873 [*DOI :* 10.1126/SCIADV.AAU1873], https://hal.inria.fr/hal-01950376

International Conferences with Proceedings

[7] F. FAGES, S. SOLIMAN.*On Robustness Computation and Optimization in BIOCHAM-4*, in "16th Int. Conf. on Computational Methods in Systems Biology", Brno, Czech Republic, September 2018, Long version with appendices [*DOI :* 10.1007/978-3-319-99429-1_18], https://hal.inria.fr/hal-01814854

Scientific Books (or Scientific Book chapters)

[8] F. FAGES, G. LE GULUDEC.*Biochemical Programs and Analog-Digital Mixed Algorithms in the Cell*, in "Life Sciences, Information Sciences", John Wiley & Sons, Inc., March 2018, p. 201-216 [*DOI :* 10.1002/9781119452713.CH19], https://hal.inria.fr/hal-01948714

Scientific Popularization

[9] F. MOLINA, J.-H. JIANG, F. FAGES.*BIOPSY: Biochemical Programming System*, November 2018, Festival Scientifique France Taiwan, ANR-MOST, Poster, https://hal.inria.fr/hal-01948702

Other Publications

[10] F. BOULIER, F. FAGES, O. RADULESCU, S. S. SAMAL, A. SCHUPPERT, W. M. SEILER, T. STURM, S. WALCHER, A. WEBER.*The SYMBIONT Project: Symbolic Methods for Biological Networks*, August 2018, F1000Research 7:1341 (poster), https://hal.inria.fr/hal-01889825

[11] V. SINGH, S. PANKAEW, M. OSTASZEWSKI, G. D. KALLIOLIAS, S. SOLIMAN, T. HELIKAR, A. NIARAKIS.*Executable Disease Networks: Adding dynamics to molecular maps*, September 2018, ECCB 2018, 17th European Conference on Computational Biology, Poster, https://hal.inria.fr/hal-01875418

Project-Team M3DISIM

Mathematical and Mechanical Modeling with Data Interaction in Simulations for Medicine

IN COLLABORATION WITH: Laboratoire de Mécanique des Solides

IN PARTNERSHIP WITH:
Ecole Polytechnique

RESEARCH CENTER
Saclay - Île-de-France

THEME
Modeling and Control for Life Sciences

Table of contents

Project-Team M3DISIM

Creation of the Team: 2013 January 01, updated into Project-Team: 2016 June 01

Keywords:

Computer Science and Digital Science:

A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.2. - Stochastic Modeling
A6.1.4. - Multiscale modeling
A6.1.5. - Multiphysics modeling
A6.2.1. - Numerical analysis of PDE and ODE
A6.3.1. - Inverse problems
A6.3.2. - Data assimilation
A6.3.4. - Model reduction
A6.4.1. - Deterministic control
A6.4.3. - Observability and Controlability
A6.4.4. - Stability and Stabilization
A6.5.1. - Solid mechanics
A6.5.2. - Fluid mechanics
A6.5.4. - Waves
A9.2. - Machine learning

Other Research Topics and Application Domains:

B1.1.3. - Developmental biology
B1.1.8. - Mathematical biology
B1.1.9. - Biomechanics and anatomy
B2.2.1. - Cardiovascular and respiratory diseases
B2.6.2. - Cardiac imaging
B2.6.3. - Biological Imaging

1. Team, Visitors, External Collaborators

Research Scientists

Dominique Chapelle [Team leader, Inria, Senior Researcher, HDR]
Philippe Moireau [Team leader, Inria, Senior Researcher, HDR]
Radomir Chabiniok [Inria, Starting Research Position]
Frédérique Clément [Inria, Senior Researcher, HDR]
Sébastien Imperiale [Inria, Researcher]

Faculty Members

Jean-Marc Allain [Ecole polytechnique, Professor, HDR]
Martin Genet [Ecole polytechnique, Associate Professor]
Patrick Le Tallec [Ecole polytechnique, Professor, HDR]

External Collaborators

Matthieu Caruel [Univ Paris-Val de Marne]
Fabrice Vallée [Assistance publique/Hôpitaux de Paris]

Hajer Methenni [CEA]

Technical Staff

Jerome Diaz [Inria, from Nov 2018]
Gautier Bureau [Inria, REO Team, until Mar 2018]

PhD Students

Ezgi Berberoglu [ETH Zurich, PhD Student, from May 2017]
Federica Caforio [Inria]
Chloe Giraudet [Ecole polytechnique, from Oct 2018]
Marija Gusseva [Inria]
Ustim Khristenko [Ecole polytechnique, until Jan 2018]
François Kimmig [Ecole polytechnique]
Arthur Le Gall [Assistance publique/Hôpitaux de Paris]
Cécile Patte [Inria]
Frédérique Robin [Inria]
Nicole Tueni [Ecole polytechnique, from Jan 2018]
Florent Wijanto [Ecole polytechnique]

Visiting Scientists

Francesco Regazzoni [Ecole polytechnique, from Sep 2018]
Katerina Solovska [Czech Technical University in Prague, Aug 2018]
Animesh Tandon [University of Texas Southwestern Medical School, Oct 2018]

Administrative Assistants

Stephanie Aubin [Inria, until Sep 2018]
Hélèna Kutniak [Inria, from Oct 2018]

2. Overall Objectives

2.1. Overall Objectives

The research carried out in the M3DISIM team has a rather global methodological perspective oriented towards biomechanics, encompassing mathematical modeling and analysis, inverse problems arising from model-data coupling, and the formulation and analysis of effective and reliable numerical procedures adapted to this overall program. We are also very keen on demonstrating the effectiveness and relevance of these methods in actual applications, usually by proof-of-concept studies carried out within various collaborations.

3. Research Program

3.1. Multi-scale modeling and coupling mechanisms for biomechanical systems, with mathematical and numerical analysis

Over the past decade, we have laid out the foundations of a multi-scale 3D model of the cardiac mechanical contraction responding to electrical activation. Several collaborations have been crucial in this enterprise, see below references. By integrating this formulation with adapted numerical methods, we are now able to represent the whole organ behavior in interaction with the blood during complete heart beats. This subject was our first achievement to combine a deep understanding of the underlying physics and physiology and our constant concern of proposing well-posed mathematical formulations and adequate numerical discretizations. In fact, we have shown that our model satisfies the essential thermo-mechanical laws, and in particular the energy balance, and proposed compatible numerical schemes that – in consequence – can be rigorously analyzed, see [6]. In the same spirit, we have formulated a poromechanical model adapted to the blood perfusion in the heart, hence precisely taking into account the large deformation of the mechanical medium, the fluid inertia and moving domain, and so that the energy balance between fluid and solid is fulfilled from the model construction to its discretization, see [7].

3.2. Inverse problems with actual data – Fundamental formulation, mathematical analysis and applications

A major challenge in the context of biomechanical modeling – and more generally in modeling for life sciences – lies in using the large amount of data available on the system to circumvent the lack of absolute modeling ground truth, since every system considered is in fact patient-specific, with possibly non-standard conditions associated with a disease. We have already developed original strategies for solving this particular type of inverse problems by adopting the observer stand-point. The idea we proposed consists in incorporating to the classical discretization of the mechanical system an estimator filter that can use the data to improve the quality of the global approximation, and concurrently identify some uncertain parameters possibly related to a diseased state of the patient. Therefore, our strategy leads to a coupled model-data system solved similarly to a usual PDE-based model, with a computational cost directly comparable to classical Galerkin approximations. We have already worked on the formulation, the mathematical and numerical analysis of the resulting system – see [5] – and the demonstration of the capabilities of this approach in the context of identification of constitutive parameters for a heart model with real data, including medical imaging, see [3].

4. Application Domains

4.1. Clinical applications

After several validation steps – based on clinical and experimental data – we have reached the point of having validated the heart model in a pre-clinical context where we have combined direct and inverse modeling in order to bring predictive answers on specific patient states. For example, we have demonstrated the predictive ability of our model to set up pacemaker devices for a specific patient in cardiac resynchronization therapies, see [11]. We have also used our parametric estimation procedure to provide a quantitative characterization of an infarct in a clinical experiment performed with pigs, see [3].

5. Highlights of the Year

5.1. Highlights of the Year

During the 8th World Congress of Biomechanics in Dublin, Martin Genet received the Young Investigator Award from the Francophone Society of Biomechanics for his talk on “A continuum relaxed growth framework for controlling growth-induced residual stresses in living tissues”.

6. New Software and Platforms

6.1. HeartLab

KEYWORDS: Computational geometry - Image analysis - Cardiac - Health - Simulation
FUNCTIONAL DESCRIPTION: The heartLab software is a library designed to perform both simulation and estimation of the heart mechanical behavior (based on various types of measurements, e.g. images).

Also included are geometric data and tools in the code to define cardiac anatomical models compatible with the simulation requirements in terms of mesh quality, fiber direction data defined within each element, and the referencing necessary for handling boundary conditions and estimation, in particular. These geometries are analytical or come from computerized tomography (CT) or magnetic resonance (MR) image data of humans or animals.

- Participants: Radomir Chabiniok, Gautier Bureau, Martin Genet, Federica Caforio, Ustim Khristenko, Dominique Chapelle and Philippe Moireau
- Contact: Philippe Moireau
- URL: https://raweb.inria.fr/rapportsactivite/RA2013/m3disim/uid14.html

6.2. Verdandi

KEYWORDS: HPC - Model - Software Components - Partial differential equation
FUNCTIONAL DESCRIPTION: Verdandi is a free and open-source (LGPL) library for data assimilation. It includes various such methods for coupling one or several numerical models and observational data. Mainly targeted at large systems arising from the discretization of partial differential equations, the library is devised as generic, which allows for applications in a wide range of problems (biology and medicine, environment, image processing, etc.). Verdandi also includes tools to ease the application of data assimilation, in particular in the management of observations or for a priori uncertainty quantification. Implemented in C++, the library may be used with models implemented in Fortran, C, C++ or Python.

- Participants: Dominique Chapelle, Gautier Bureau, Nicolas Claude, Philippe Moireau and Vivien Mallet
- Contact: Vivien Mallet
- URL: http://verdandi.gforge.inria.fr/

6.3. CardiacLab

KEYWORDS: Cardiovascular and respiratory systems - Matlab - Real time
FUNCTIONAL DESCRIPTION: CardiacLab is a MATLAB toolbox allowing to perform "real-time" cardiac simulations using 0D models of the cardiovascular systems. Its modular development includes (1) a module integrating the mechanical dynamics of the cavity taking into account its particular geometry, (2) a module allowing to choose a micro-model of the cardiac contraction, (3) a module of phase management, (4) a circulation module based on Windkessel models or more advanced 1D flows models, and (5) a perfusion module. The objective of this code is threefold: (1) demonstrate to students, engineers, medical doctors, the interest of modeling in cardiac applications, (2) unify our original modeling developments with the possibility to evaluate them with previous team developments before integrating them into 3D complex formulations, and (3) explore some avenues pertaining to real-time simulat

- Participants: Sebastien Impériale, Martin Genet, Federica Caforio, Ustim Khristenko, Peter Baumgartner, Radomir Chabiniok, François Kimmig and Arthur Le Gall
- Contact: Philippe Moireau
- URL: https://gitlab.inria.fr/M3DISIM/CardiacLab

6.4. MoReFEM

Modeling Research with the Finite Element Method
KEYWORDS: HPC - Multiphysics modelling - Data assimilation
FUNCTIONAL DESCRIPTION: MoReFEM is a HPC finite element library for simulating multiphysics evolution problems like the ones encounter in cardiac modeling (electrophysiology, structure and fluid mechanics, transport-diffusion, wave equations)

- Participants: Philippe Moireau, Patrick Le Tallec, Antoine Olivier, Dominique Chapelle, Ustim Khristenko, François Kimmig, Gautier Bureau and Sébastien Gilles
- Contact: Sébastien Gilles
- URL: https://gitlab.inria.fr/MoReFEM

7. New Results

7.1. Mathematical and Mechanical Modeling

7.1.1. Microscopic model of collagen fiber

Participants: Florent Wijanto, Matthieu Caruel, Jean-Marc Allain [correspondant].

Our studies on collagen tissues have shown that the collagen fibers are able to elongate inelastically under stretch. In tendon, this effect has been attributed to the non-permanent cross-bridges which connect the different collagen fibrils (to assemble a fiber). This sliding effect appears experimentally to be reversible (at least partially) if the tissue is left long enough at its initial resting length. However, this sliding is classically included as an irreversible plastic response, or as a damage of the tissue. We are building a model based on a stochastic description of the binding and unbinding of the cross-bridges. This approach will enable us to have a microscopically based picture of the sliding, which will be able to explain some alteration in case of ageing or of pathological alterations of the tissue. At the moment, we have shown the importance of the density of cross-bridges in the cooperative response of the system. A publication is in preparation on the topic.

7.1.2. Stochastic modeling of chemical-mechanical coupling in striated muscles

Participants: Matthieu Caruel, Dominique Chapelle [correspondant], Philippe Moireau.

We propose a chemical-mechanical model of myosin heads in sarcomeres, within the classical description of rigid sliding filaments. In our case, myosin heads have two mechanical degrees-of-freedom (dofs) – one of which associated with the so-called power stroke – and two possible chemical states, i.e. bound to an actin site or not. Our major motivations are twofold: (1) to derive a multiscale coupled chemical-mechanical model, and (2) to thus account – at the macroscopic scale – for mechanical phenomena that are out of reach for classical muscle models. This model is first written in the form of Langevin stochastic equations, and we are then able to obtain the corresponding Fokker-Planck partial differential equations governing the probability density functions associated with the mechanical dofs and chemical states. This second form is important, as it allows to monitor muscle energetics, and also to compare our model with classical ones, such as the Huxley'57 model to which our equations are shown to reduce under two different types of simplifying assumptions. This provides insight, and gives a Langevin form for Huxley'57. We then show how we can calibrate our model based on experimental data – taken here for skeletal muscles – and numerical simulations demonstrate the adequacy of the model to represent complex physiological phenomena, in particular the fast isometric transients in which the power stroke is known to have a crucial role, thus circumventing a limitation of many classical models. This work is accepted for publication in BMMB.

7.1.3. The importance of the pericardium for cardiac biomechanics

Participant: Radomir Chabiniok [correspondant].

The human heart is enclosed in the pericardial cavity. The pericardium consists of a layered thin sac and is separated from the myocardium by a thin film of fluid. It provides a fixture in space and frictionless sliding of the myocardium. The influence of the pericardium is essential for predictive mechanical simulations of the heart. However, there is no consensus on physiologically correct and computationally tractable pericardial boundary conditions. Here we propose to model the pericardial influence as a parallel spring and dashpot acting in normal direction to the epicardium. Using a four-chamber geometry, we compare a model with pericardial boundary conditions to a model with fixated apex. The influence of pericardial stiffness is demonstrated in a parametric study. Comparing simulation results to measurements from cine magnetic resonance imaging reveals that adding pericardial boundary conditions yields a better approximation with respect to atrioventricular plane displacement, atrial filling, and overall spatial approximation error. We demonstrate that this simple model of pericardial-myocardial interaction can correctly predict the pumping mechanisms of the heart as previously assessed in clinical studies. Utilizing a pericardial model can not only provide much more realistic cardiac mechanics simulations but also allows new insights into pericardial-myocardial interaction which cannot be assessed in clinical measurements yet. The work was accepted for publication in Biomechanics and Modeling in Mechanobiology [26], and is a joint work with Technical University in Munich, Germany (group of W.A. Wall) and Bernoulli Institute for Mathematics at University of Groningen, The Netherlands (C. Bertoglio).

7.1.4. *Solving 2D linear isotropic elastodynamics by means of scalar potentials: a new challenge for finite elements*

Participants: Sébastien Imperiale, Patrick Joly [Poems].

In this work we present a method for the computation of numerical solutions of 2D homogeneous isotropic elastodynamics equations by solving scalar wave equations. These equations act on the potentials of a Helmholtz decomposition of the displacement field and are decoupled inside the propagation domain. We detail how these equations are coupled at the boundary depending on the nature of the boundary condition satisfied by the displacement field. After presenting the case of rigid boundary conditions, that presents no specific difficulty, we tackle the challenging case of free surface boundary conditions that presents severe stability issues if a straightforward approach is used. We introduce an adequate functional framework as well as a time domain mixed formulation to circumvent these issues. Numerical results confirm the stability of the proposed approach.

7.1.5. *Lung multiscale poromechanical modeling, from breathing to pulmonary fibrosis-induced chronic remodeling*

Participants: Cécile Patte [correspondant], Martin Genet, Dominique Chapelle.

Pulmonary diseases are about to become the third cause of death in the world. One on them, Idiopathic Pulmonary Fibrosis (IPF), which involves thickening, stiffening and destruction the alveolar walls, remains poorly understood, diagnosed and treated. It has been hypothesized, however, that IPF involves a mechanical vicious circle, where fibrosis induces higher stresses, which in turns favors fibrosis. In this project, we intend to better understand the role of mechanics in the disease progression, in order to improve diagnosis and prognosis. We model the lung behavior during breathing at organ-scale, based on a poromechanical theory, previously established in the team. Then we estimate the regional mechanical properties of the lung, based on clinical data. In the future, the procedure can be used as a prognostic tool by the clinicians.

7.1.6. *Mathematical modelling of transient shear wave elastography in the heart*

Participants: Federica Caforio [correspondant], Sébastien Imperiale.

The aim of this work is to provide a mathematical model of the excitation and the resulting shear wave propagation in Acoustic Radiation Force (ARF)-based shear wave cardiac elastography. Our approach is based on asymptotic analysis; more precisely, it consists in considering a family of problems, parametrised by a small parameter inversely proportional to the excitation frequency of the probes, the viscosity and the velocity of pressure wave propagation. We derive a simplified model for the expression of the ARF by investigating the limit behaviour of the solution when the small parameter goes to zero. By formal asymptotic analysis - an asymptotic expansion of the solution is used - we show that the leading order term of the expansion is the underlying nonlinear cardiac mechanics. Subsequently, two corrector terms are computed. The first is a fast-oscillating pressure wave generated by the probes, solution of a Helmholtz equation at every time instant. The second corrector term consists in an elastic field with prescribed divergence, having a function of the first corrector as a source term. This field corresponds to the shear acoustic wave induced by the ARF. We also confirm that, in cardiac mechanics, the presence of viscosity in the model is essential to derive an expression of the shear wave propagation from the ARF, and that this phenomenon is related to the nonlinearity of the partial differential equation.

7.1.7. *Analysis and calibration of a linear model for structured cell populations with unidirectional motion : Application to the morphogenesis of ovarian follicles*

Participants: Frédérique Clément, Frédérique Robin [correspondant], Romain Yvinec [INRA].

In [41], we have analyzed a multi-type age dependent model for cell populations subject to unidirectional motion, in both a stochastic and deterministic framework. Cells are distributed into successive layers; they may divide and move irreversibly from one layer to the next. We have adapted results on the large-time convergence of PDE systems and branching processes to our context, where the Perron-Frobenius or Krein-Rutman theorems cannot be applied. We have derived explicit analytical formulas for the asymptotic cell number moments, and the stable age distribution. We have illustrated these results numerically and we have applied them to the study of the morphodynamics of ovarian follicles. We have proven the structural parameter identifiability of our model in the case of age independent division rates. Using a set of experimental biological data, we have estimated the model parameters to fit the changes in the cell numbers in each layer during the early stages of follicle development.

7.1.8. A multiscale mathematical model of cell dynamics during neurogenesis in the mouse cerebral cortex

Participants: Frédérique Clément [correspondant], Marie Postel [Sorbonne Universités].

Work in collaboration with Sylvie Schneider-Maunoury (Sorbonne Universités), Alice Karam (Sorbonne Universités), Guillaume Pézeron (MNHN). Neurogenesis in the murine cerebral cortex involves the coordinated divisions of two main types of progenitor cells, whose numbers, division modes and cell cycle durations set up the final neuronal output. To understand the respective roles of these factors in the neurogenesis process, we have combined experimental in vivo studies with mathematical modeling and numerical simulations of the dynamics of neural progenitor cells [43]. A special focus has been put on the population of intermediate progenitors (IPs), a transit amplifying progenitor type critically involved in the size of the final neuron pool. Our multiscale formalism describing IP dynamics allows one to track the progression of cells along the subsequent phases of the cell cycle, as well as the temporal evolution of the different cell numbers. Our model takes into account the dividing apical progenitors (AP) engaged into neurogenesis, both neurogenic and proliferative IPs, and the newborn neurons. The transfer rates from one population to another are subject to the mode of division (symmetric, asymmetric, neurogenic) and may be time-varying. The model outputs have been successfully fitted to experimental cell numbers from mouse embryos at different stages of cortical development, taking into account IPs and neurons, in order to adjust the numerical parameters. Applying the model to a mouse mutant for Ftm/Rpgrip1l, a gene involved in human ciliopathies with severe brain abnormalities, has revealed a shortening of the neurogenic period associated with an increased influx of newborn IPs from apical progenitors at mid-neurogenesis. Additional information have been provided on cell kinetics, such as the mitotic and S phase indexes, and neurogenic fraction. Our model can be used to study other mouse mutants with cortical neurogenesis defects and can be adapted to study the importance of progenitor dynamics in cortical evolution and human diseases.

7.1.9. Advances in computational modeling approaches of pituitary gonadotropin signaling

Participants: Frédérique Clément [correspondant], Romain Yvinec [INRA].

Work in collaboration with Pascale Crépieux, Anne Poupon and Éric Reiter (INRA). We have reviewed thoroughly the state-of-the-art in computational modeling approaches of pituitary gonadotropin signaling [30]. Pituitary gonadotropins play an essential and pivotal role in the control of human and animal reproduction within the hypothalamic-pituitary-gonadal (HPG) axis. The computational modeling of pituitary gonadotropin signaling encompasses phenomena of different natures such as the dynamic encoding of gonadotropin secretion, and the intracellular cascades triggered by gonadotropin binding to their cognate receptors, resulting in a variety of biological outcomes. We have overviewed historical and ongoing issues in modeling and data analysis related to gonadotropin secretion in the field of both physiology and neuro-endocrinology. We have mentioned the different mathematical formalisms involved, their interest and limits. We have discussed open statistical questions in signal analysis associated with key endocrine issues. We have also reviewed recent advances in the modeling of the intracellular pathways activated by gonadotropins, which yields promising development for innovative approaches in drug discovery. The greatest challenge to be tackled in computational modeling of pituitary gonadotropin signaling is the embedding of gonadotropin signaling within its natural multiscale environment, from the single cell level, to the organic and whole HPG level.

The development of modeling approaches of G protein-coupled receptor signaling, together with multicellular systems biology may lead to unexampled mechanistic understanding with critical expected fallouts in the therapeutic management of reproduction.

7.1.10. Structured cell population dynamics applied to the early development of ovarian follicles

Participants: Frédérique Clément, Frédérique Robin [correspondant], Romain Yvinec [INRA].

The ovarian follicles are the basic anatomical and functional units of the ovaries, which are renewed from a quiescent pool all along reproductive life. Follicular development involves a finely tuned sequence of growth and maturation processes, involving complex cell dynamics. Understanding follicular development is a crucial issue for the management of reproduction in a clinical or breeding context, and for the preservation of endangered species. In their early stages of development, ovarian follicles are made up of a germ cell (oocyte), whose diameter increases steadily, and of surrounding proliferating somatic cells, which are layered in a globally spherical and compact structure. We have designed a modeling approach dedicated to the initiation phase of follicle development. The initiation phase is described by joint stochastic dynamics accounting for cell shape transitions (from a flattened to a cuboidal shape) and proliferation of reshaped cells. We have then derived the mean time elapsed before all cells have changed shapes and the corresponding increment in the total cell number, which is fitted to experimental data retrieved from primordial follicles (single layered follicle with only flattened cells) and primary follicles (single layered follicles with only cuboidal cells).

7.1.11. Newton-Krylov method for computing the cyclic steady states of evolution problems in non-linear mechanics

Participants: Ustim Khristenko, Patrick Le Tallec [correspondant].

This work is focused on the Newton-Krylov technique for computing the steady cyclic states of evolution problems in nonlinear mechanics with space-time periodicity conditions. This kind of problems can be faced, for instance, in the modeling of a rolling tire with a periodic tread pattern, where the cyclic state satisfies “rolling” periodicity condition, including shifts both in time and space. The Newton-Krylov method is a combination of a Newton nonlinear solver with a Krylov linear solver, looking for the initial state, which provides the space-time periodic solution. The convergence of the Krylov iterations is proved to hold in presence of an adequate preconditioner. After preconditioning, the Newton-Krylov method can be also considered as an observer-controller method, correcting the transient solution of the initial value problem after each period. Using information stored while computing the residual, the Krylov solver computation time becomes negligible with respect to the residual computation time. The method has been analyzed and tested on academic applications and compared with the standard evolution (fixed point) method. Finally, it has been implemented into the Michelin industrial code, applied to a full 3D rolling tire model.

7.1.12. Delayed feedback control method for computing the cyclic steady states of evolution problems

Participants: Ustim Khristenko, Patrick Le Tallec [correspondant].

This work is focused on fast techniques for computing the cyclic steady states of evolution problems in non-linear mechanics with space–time periodicity conditions. In industrial applications, in order to avoid the inversion of very large matrices, such a cyclic solution is usually computed as an asymptotic limit of the associated initial value problem with arbitrary initial data. However, when the relaxation time is high, convergence to the limit cycle can be very slow. In such cases nonetheless, one is not interested in the transient solution, but only in a fast access to the limit cycle. Thus, in this work we modify the problem, introducing the time-delayed feedback control, which is widely used for stabilization of unstable periodic orbits. In our framework it is applied to an initially stable system in order to accelerate its convergence to the limit cycle. Moreover, the control term, based on the space–time periodicity error, includes both shifts in time and in space. Our main result is the optimal form of the control term for a very general class of linear evolution problems, providing the fastest convergence to the cyclic solution, which has been further extended and studied in the non-linear case. Efficiency of the method increases with the problem's relaxation time. The method has been tested using academic applications and compared to the non-controlled asymptotic convergence as well as to the Newton–Krylov shooting algorithm. Finally, the method has been implemented into the Michelin industrial code, applied to a full 3D rolling tyre model.

7.2. Numerical Methods

7.2.1. Numerical analysis for an energy-preserving total discretization of a poromechanics model with inf-sup stability

Participants: Dominique Chapelle [correspondant], Philippe Moireau.

We consider a previously proposed general nonlinear poromechanical formulation, and we derive a linearized version of this model. For this linearized model, we obtain an existence result and we propose a complete discretization strategy – in time and space – with a special concern for issues associated with incompressible or nearly-incompressible behavior. We provide a detailed mathematical analysis of this strategy, the main result being an error estimate uniform with respect to the compressibility parameter. We then illustrate our approach with detailed simulation results and we numerically investigate the importance of the assumptions made in the analysis, including the fulfillment of specific inf-sup conditions. This work is accepted for publication in Acta Mathematicae Applicatae Sinica.

7.2.2. Efficient estimation of personalized biventricular mechanical function employing gradient-based optimization

Participant: Martin Genet [correspondant].

Individually personalized computational models of heart mechanics can be used to estimate important physiological and clinically-relevant quantities that are difficult, if not impossible, to directly measure in the beating heart. Here, we present a novel and efficient framework for creating patient-specific biventricular models using a gradient-based data assimilation method for evaluating regional myocardial contractility and estimating myofiber stress. These simulations can be performed on a regular laptop in less than 2 hours and produce excellent fit between measured and simulated volume and strain data through the entire cardiac cycle. By applying the framework using data obtained from 3 healthy human biventricles, we extracted clinically important quantities as well as explored the role of fiber angles on heart function. Our results show that steep fiber angles at the endocardium and epicardium are required to produce simulated motion compatible with measured strain and volume data. We also find that the contraction and subsequent systolic stresses in the right ventricle are significantly lower than that in the left ventricle. Variability of the estimated quantities with respect to both patient data and modeling choices are also found to be low. Because of its high efficiency, this framework may be applicable to modeling of patient specific cardiac mechanics for diagnostic purposes.

7.2.3. Equilibrated warping: Finite element image registration with finite strain equilibrium gap regularization

Participant: Martin Genet [correspondant].

In this work, we propose a novel continuum finite strain formulation of the equilibrium gap regularization for image registration. The equilibrium gap regularization essentially penalizes any deviation from the solution of a hyperelastic body in equilibrium with arbitrary loads prescribed at the boundary. It thus represents a regularization with strong mechanical basis, especially suited for cardiac image analysis. We describe the consistent linearization and discretization of the regularized image registration problem, in the framework of the finite elements method. The method is implemented using FEniCS & VTK, and distributed as a freely available python library. We show that the equilibrated warping method is effective and robust: regularization strength and image noise have minimal impact on motion tracking, especially when compared to strain-based regularization methods such as hyperelastic warping. We also show that equilibrated warping is able to extract main deformation features on both tagged and untagged cardiac magnetic resonance images.

7.2.4. Thermodynamic properties of muscle contraction models and associated discrete-time principles

Participants: François Kimmig [correspondant], Dominique Chapelle, Philippe Moireau.

Considering a large class of muscle contraction models accounting for actin-myosin interaction, we present a mathematical setting in which solution properties can be established, including fundamental thermodynamic balances. Moreover, we propose a complete discretization strategy for which we are also able to obtain discrete versions of the thermodynamic balances and other properties. Our major objective is to show how the thermodynamics of such models can be tracked after discretization, including when they are coupled to a macroscopic muscle formulation in the realm of continuum mechanics. Our approach allows to carefully identify the sources of energy and entropy in the system, and to follow them up to the numerical applications.

7.2.5. A conservative penalisation strategy for the semi-implicit time discretisation of the incompressible elastodynamics equation

Participants: Federica Caforio [correspondant], Sébastien Imperiale.

The principal aim of this work is to provide an adapted numerical scheme for the approximation of elastic wave propagation in incompressible solids. We rely on high-order conforming finite element with mass lumping for space discretisation and implicit/explicit, second-order, energy-preserving time discretisation. The time step restriction only depends on the shear wave velocity and at each time step a Poisson problem must be solved to account for the incompressibility constraint that is imposed by penalisation techniques.

7.2.6. High-order discrete Fourier transform for the solution of the poisson equation

Participants: Federica Caforio [correspondant], Sébastien Imperiale.

The aim of this work is to propose a novel, fast, matrix-free solver for the Poisson problem discretised with High-Order Spectral Element Methods (HO-SEM). This method is based on the use of the Discrete Fourier Transform to reduce the problem to the inversion of the symbol of the operator in frequency space. The solver proposed is endowed with several properties. First, it preserves the efficiency of standard FFT algorithm; then, the matrix storage is minimised; a pseudo- explicit Singular Value Decomposition (SVD) is used for the inversion of the symbols; finally, it can be easily extended to multiple dimensions and non-periodic boundary conditions. In particular, due to the underlying HO-SEM discretisation, the multi-dimensional symbol of the operator can be efficiently computed from the one-dimensional symbol by tensorisation.

7.3. Inverse Problems

7.3.1. Analysis of an observer strategy for initial state reconstruction of wave-like systems in unbounded domain

Participants: Sébastien Imperiale, Philippe Moireau [correspondant].

We are interested in reconstructing the initial condition of a wave equation in an unbounded domain configuration from measurements available in time on a subdomain. To solve this problem, we adopt an iterative strategy of reconstruction based on observers and time reversal adjoint formulations. We prove the convergence of our reconstruction algorithm with perfect measurements and its robustness to noise. Moreover, we develop a complete strategy to practically solve this problem on a bounded domain using artificial transparent boundary conditions to account for the exterior domain. Our work then demonstrates that the consistency error introduced by the use of approximate transparent boundary conditions is compensated by the stabilization properties obtained from the use of the available measurements, hence allowing to still be able to reconstruct the unknown initial condition. This work is accepted with minor revision for publication in COCV.

7.4. Experimental Assessments

7.4.1. Mathematical modeling and experimental validation of flow through aortic valve

Participant: Radomir Chabiniok [correspondant].

Assessment of the valvular diseases by phase-contrast magnetic resonance imaging (MRI) has known limits due to limited spatial-temporal resolution of MRI and artifacts intrinsic to the method. This problem is addressed by the collaborative work of the Institute for Clinical and Experimental Medicine in Prague (IKEM, participants J. Tintera and R. Galabov) and the mathematical modeling group at the Czech Technical University in Prague (CTU, participants P. Paus, R. Fucik), additionally with the combined clinical cardiovascular MRI & modeling expertise of R. Chabiniok (Inria). A flow phantom was constructed at IKEM and used to perform an extensive experimental study targeted to capture the phenomena in valvular stenosis / regurgitation. The Mathematical modeling group at CTU then performed flow simulations by using the techniques of Lattice-Boltzmann method and their high-performance computing GPU implementations. This work is shedding light into possibly significant factors limiting the direct interpretation of PC MRI and opening the way into interaction of PC MRI data with mathematical model as a "smart filtering" of flow exam.

7.4.2. Skin multiscale mechanics

Participant: Jean-Marc Allain [correspondant].

Skin is a complex, multi-layered organ, with important functions in the protection of the body. The dermis provides structural support to the epidermal barrier, and thus has attracted a large number of mechanical studies. As the dermis is made of a mixture of stiff fibres embedded in a soft non-fibrillar matrix, it is classically considered that its mechanical response is based on an initial alignment of the fibres, followed by the stretching of the aligned fibres. Using a recently developed set-up combining multiphoton microscopy with mechanical assay, we imaged the fibres network evolution during dermis stretching. These observations, combined with a wide set of mechanical tests, allowed us to challenge the classical microstructural interpretation of the mechanical properties of the dermis: we observed a continuous alignment of the collagen fibres along the stretching. All our results can be explained if each fibre contributes by a given stress to the global response. This plastic response is likely due to inner sliding inside each fibre. The non-linear mechanical response is due to structural effects of the fibres network in interaction with the surrounding non-linear matrix. This multiscale interpretation explains our results on genetically-modified mice with a simple alteration of the dermis microstructure. Our previous works have led us to write this year one review article and one chapter of book on multiscale skin biomechanics, to be published next year.

7.4.3. Cornea biomechanics

Participants: Chloé Giraudet, Jean-Marc Allain [correspondant], Patrick Le Tallec.

Cornea is the outer part of the eye. It is a curved transparent organ, which gives 2/3 of the focalisation capacity of the eye. Microscopically, it is made mostly of collagen fibres (as skin) organised in cristal-like lamellae of few micrometers of height and a hundred micrometers in length and width. The lamellae are piled up in a plywood structure, creating a millimetre-thick tissue. Between the lamellae, some cells are present to repair and regenerate the tissue. However, this simple image of the organisation of the collagen is in fact too simple and a more complex heterogeneous organisation has been recently described, with in particular some striae (called the Vogt striae). In C. Giraudet's PhD, we propose to explore the link between microstructure organisation of the collagen in the cornea and mechanical properties. To do so, we will first start by proposing an extension of classical mechanical models (such as Holzapfel's law or others) to the specific case of the cornea. This model will be tested against mechanical assays made under advanced optical microscopes to test first if the model can correctly predict the strain field in volume, and secondly if it correctly predicts the evolution of the lamellae microstructure at different stretch levels. At the moment, we have developed the tools to mechanically test the cornea, but also to build a finite element simulation using the real shape of the cornea we are looking at.

7.4.4. Multiscale properties of the passive cardiac muscle

Participants: Nicole Tueni, Jean-Marc Allain [correspondant], Martin Genet.

We are interested in understanding the effect of the remodelling of cardiac tissues after a disease. Cardiac tissues are mostly made of muscle cells. They can remodel themselves in response to an alteration of their normal response by modifying the sizes and the geometries of the cells in the tissue. Nowadays, we are able to describe the active and passive response of a cardiac tissue, assuming we know the main orientation of the cells inside. However, we do not have models which include explicitly the microstructural cellular organization. Such complex models will be strongly beneficial to determine the consequences of local alterations of the muscle behaviour. In N. Tueni's PhD, we are investigating this multi-scale relationship. To do so, we are imaging the organization at the microscale, while measuring the mechanical properties. These results will be the building block to test and develop mechanical models of the cardiac tissues.

7.4.5. Mechano-perception at the cell level

Participant: Jean-Marc Allain [correspondant].

All cells and organisms experience mechanical forces. Plants along their life are submitted from their environment to long lasting sustained stresses and to recurrent cyclic loading/unloading due to wind or water stream. Mechanical stimulations induce short-term cellular responses, leading to mechanoresponsive gene activation followed by long-term responses permitting structural reinforcement at the whole-plant level. We show that the Mechanosensitive channel Small conductance-Like 10 (MSL10) contributes to oscillation perception at the cell level. This channel responds to pulsed membrane stretching with rapid activation and relaxation. Furthermore, oscillatory pressure stimulation modulates its activity, with increased open probability upon oscillatory than during sustained stimulation. Combined with the adequate localization of MSL10 in plant shoot and leaves, its ability to detect oscillatory deformation at the molecular-scale is relevant for a function of this channel in oscillatory perception in plant.

7.5. Clinical Applications

7.5.1. Exploring kinetic energy as a new marker of cardiac function in the single ventricle circulation

Participants: Radomir Chabiniok [correspondant], Tarique Hussain [ToFMOD].

Ventricular volumetric ejection fraction (VV EF) is often normal in patients with single ventricle circulations despite them experiencing symptoms related to circulatory failure. We sought to determine if kinetic energy (KE) could be a better marker of ventricular performance. KE was prospectively quantified using four-dimensional flow MRI in 41 patients with a single ventricle circulation (aged 0.5-28 yr) and compared with 43 healthy volunteers (aged 1.5-62 yr) and 14 patients with left ventricular (LV) dysfunction (aged 28-79 yr). Intraventricular end-diastolic blood was tracked through systole and divided into ejected and residual blood components. Two ejection fraction (EF) metrics were devised based on the KE of the ejected component over the total of both the ejected and residual components using 1) instantaneous peak KE to assess KE EF or 2) summating individual peak particle energy (PE) to assess PE EF. KE metrics are markers of healthy cardiac function. PE EF may be useful in grading dysfunction. The work was published in Journal of Applied Physiology (J Appl Physiol 125: 889-900, 2018), [29]. The work represents a collaboration with King's College London (J. Wong, K. Pushparajah, R. Razavi) and with UT Southwestern Dallas (T. Hussain, the member of Inria Associate team ToFMOD).

7.5.2. Using a patient-specific biomechanical cardiovascular model to estimate continuously Left Ventricular Pressure Volume Loop: A proof of concept study

Participants: Arthur Le Gall, Fabrice Vallée, Philippe Moireau, Dominique Chapelle, Radomir Chabiniok [correspondant].

Pressure Volume loops (PV loops) could contribute to optimise haemodynamic managements. While the invasiveness of PV loop acquisition prevents it from being routinely used during surgery, cardiovascular modelling could represent an alternative. Using continuous recording of aortic pressure and flow, we aimed at calibrating a patient-specific model and at interpreting the simulated PV loop during administration of noradrenaline (NOR). This study is the first to allow continuous PV loop monitoring during general anaesthesia. The work was pursued in the collaboration with Lariboisiere Hospital in Paris (A. Le Gall and F. Vallée, both dually affiliated at Inria and at AP-HP, “poste d’accueil”).

7.5.3. *Augmenting the interpretation of cardiac MRI by biomechanical modeling: Application to Tetralogy of Fallot*

Participants: Marija Gusseva, Philippe Moireau, Tarique Hussain [ToFMod], Gerald Greil [ToFMod], Animesh Tandon [ToFMOD], Dominique Chapelle, Radomir Chabiniok [correspondant].

The particularity of the mixed-valve disease – pulmonary regurgitation often combined with a stenosis – requested to extend our model-representation of the valve to allow the backflow during the heart relaxation. For each patient, biomechanical models of their left and right ventricles (LV, RV) were set up. These models then allowed to investigate the functional properties of dilated right ventricles (RV) with incompetent pulmonary valves and of the pulmonary circulation, properties not directly visible in the clinical data. In particular, immediately after deploying the new valve we could observe a decrease of RV contractility by 15%, while the output of RV into pulmonary circulation has increased. This suggests a positive immediate outcome, as the energy needs for function of RV will decrease. The higher cardiac output also suggests an increase of the filling of LV (preload), which could contribute to an improvement of LV function. The model also uncovered a decrease of resistance in the pulmonary circulation. This very preliminary result might suggest some pathophysiological changes, which are typically not thought of in clinics.

This work is pursued under the objectives of the Inria Associate Team ToFMod (T. Hussain, G. Greil, A. Tandon are members of ToFMOD and affiliated at UT Southwestern Medical Center Dallas, USA), the work was accepted for a conference of International Society of Magnetic Resonance in Medicine 2018 and is in preparation for publication.

7.5.4. *Longitudinal study of ventricular remodeling and reverse-remodeling in tetralogy of Fallot patients using CMR coupled with biomechanical modelling*

Participants: Marija Gusseva, Tarique Hussain [ToFMod], Animesh Tandon [ToFMod], Dominique Chapelle, Radomir Chabiniok [correspondant].

A preliminary study was performed with the patient-specific models for RV and pulmonary circulations set up from three datasets including the 6-months post-PVR follow-up exams obtained in late 2018 from King’s College London. Clinical data analyses show a positive result of pulmonary replacement therapy (PVR) and normalization of the RV size, i.e. the so-called reverse-remodeling of the pathologically dilated RV, in all three patients. The biomechanical modeling suggests a further reduction of the active stress needed to be developed by RV (contractility), i.e. a long-term unloading of the previously overloaded ventricle.

This work is pursued under the objectives of the Inria Associate Team ToFMOD (T. Hussain, A. Tandon are members of ToFMOD and affiliated at UT Southwestern Medical Center Dallas, USA). The main partner in this task is King’s College London (“Other Participant” in the ToFMOD Associate team, K. Pushparajah, M. Jones, S. Qureshi) who provided unique clinical data of patients with a long-term follow-up after PVR. The work was submitted to the conference of International Society of Magnetic Resonance in Medicine 2019 – the world-wide major scientific & clinical event when MR data are involved.

7.5.5. *Optical flow-based non-rigid registration of cardiac MR images*

Participant: Radomir Chabiniok [correspondant].

This work deals with non-rigid registration of cardiac MR images, particularly the MOLLI sequences. MOLLI sequence consists of 11 heart images acquired over 17 cardiac cycles. The images of MOLLI sequence are used for pixel-wise estimation of T1 relaxation time values. In this case the registration is necessary to correct the deformations that occur because of the patient’s imperfect breath-holding during the acquisition. The main characteristics of the MOLLI sequence is the evolving intensity of the tissues and also large variations of the image contrast. This characteristics of the sequence make the registration process challenging and make the use of intensity-based registration method impossible. For this purpose, we propose a method based on optical flow, using information obtained by image segmentation. The first step of the registration process, is segmentation of the regions of interest, using the level set method. The segmented objects are represented by distance maps. The transformation between original images is determined by applying the optical flow method to the distance maps. The registration process is independent of the varying intensity and takes into account only the shape and position of the segmented areas, such as the myocardium or the ventricles. The implementation of the proposed method is described and the method is tested on several MOLLI sequences. The results are compared to the results of methods based on maximisation of mutual information, and the proposed method performs better for the images with significant changes in intensity.

The work represents a collaborative project with Institute for Clinical and Experimental Medicine (IKEM) Prague (J. Tintera) and with Czech Technical University in Prague (K. Solovska, T. Oberhuber).

7.5.6. *Quantification of biventricular strains in heart failure with preserved ejection fraction using hyperelastic warping method*

Participant: Martin Genet.

Heart failure (HF) imposes a major global health care burden on society and suffering on the individual. About 50% of HF patients have preserved ejection fraction (HFpEF). More intricate and comprehensive measurement-focused imaging of multiple strain components may aid in the diagnosis and elucidation of this disease. Here, we describe the development of a semi-automated hyperelastic warping method for rapid comprehensive assessment of biventricular circumferential, longitudinal, and radial strains that is physiological meaningful and reproducible. We recruited and performed cardiac magnetic resonance (CMR) imaging on 30 subjects [10 HFpEF, 10 HF with reduced ejection fraction patients (HFrEF) and 10 healthy controls]. In each subject, a three-dimensional heart model including left ventricle (LV), right ventricle (RV), and septum was reconstructed from CMR images. The hyperelastic warping method was used to reference the segmented model with the target images and biventricular circumferential, longitudinal, and radial strain–time curves were obtained. The peak systolic strains are then measured and analyzed in this study. The ROC analysis indicated LV peak systolic circumferential strain to be the most sensitive marker for differentiating HFpEF from healthy controls. Our results suggest that the hyperelastic warping method with the CMR-derived strains may reveal subtle impairment in HF biventricular mechanics, in particular despite a “normal” ventricular ejection fraction in HFpEF.

7.5.7. *Extra corporeal life support for cardiac arrest patients with post-cardiac arrest syndrome: the ECCAR study*

Participant: Arthur Le Gall.

Purpose: Post-Cardiac Arrest Shock (PCAS) occurring after resuscitated cardiac arrest (CA), is a main cause of early death. Extra-Corporeal Life Support (ECLS) could be useful pending recovery of myocardial failure. We aimed to describe our PCAS population, and factors associated with ECLS initiation. Materials and Methods: This analysis included 924 patients admitted in two intensive care units (ICU) between 2005 and 2014 for CA and PCAS, and, of those patients, 43 patients for whom an ECLS was initiated. Neurological and ECLS-related outcomes were gathered retrospectively. Conclusions: ECLS, as a salvage therapy for PCAS, could represent an acceptable alternative for highly selected patients.

7.5.8. *Evaluation of cardiac output variations with the peripheral pulse pressure to mean arterial pressure ratio.*

Participant: Arthur Le Gall.

Cardiac output (CO) optimisation during surgery reduces post-operative morbidity. Various methods based on pulse pressure analysis have been developed to overcome difficulties to measure accurate CO variations in standard anaesthetic settings. Several of these methods include, among other parameters, the ratio of pulse pressure to mean arterial pressure (PP/MAP). The aim of this study was to evaluate whether the ratio of radial pulse pressure to mean arterial pressure (ΔPPrad/MAP) could track CO variations (Δ CO) induced by various therapeutic interventions such as fluid infusions and vasopressors boluses [phenylephrine (PE), norepinephrine (NA) or ephedrine (EP)] in the operating room. Trans-oesophageal Doppler signal and pressure waveforms were recorded in patients undergoing neurosurgery. CO and PPrad/MAP were recorded before and after fluid challenges, PE, NA and EP bolus infusions as medically required during their anaesthesia. ΔPPrad/MAP tracked ΔCO variations during PE and NA vasopressor challenges. However, after positive fluid challenge or EP boluses, ΔPPrad/MAP was not as performant to track ΔCO which could make the use of this ratio difficult in current clinical practice.

7.5.9. Perioperative management of patients with coronary artery disease undergoing non-cardiac surgery: Summary from the French Society of Anaesthesia and Intensive Care Medicine 2017 convention

Participant: Arthur Le Gall.

This review summarises the specific stakes of preoperative, intraoperative, and postoperative periods of patients with coronary artery disease undergoing non-cardiac surgery. All practitioners involved in the perioperative management of such high cardiac risk patients should be aware of the modern concepts expected to decrease major adverse cardiac events and improve short- and long-term outcomes. A multidisciplinary approach via a functional heart team including anaesthesiologists, cardiologists and surgeons must be encouraged. Rational and algorithm-guided management of those patients should be known and implemented from preoperative to postoperative period.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- Contract with start-up 3c-industry for quantitative imaging of their printed product (1.5keuros)
- Contract with L'Oreal for the development of an experimental set-up (29.8keuros)

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. Other funding

IPM-MS project (for Imagerie Polarimétrique de Mueller pour la réalisation d'un système original de caractérisation des propriétés mécaniques des Matériaux Structurés). 50k€ funded by the LABEX Lasips. This project, which involves the LPICM laboratory (Ecole Polytechnique, CNRS), the LMS (Ecole Polytechnique, CNRS, Mines ParisTech) and the Centre des Matériaux (Mines ParisTech), aims at developing an optical tool to study the link between the mechanical properties of a material and its hierarchical organization. Despite the development of new methods to observe the microstructure, one of the limitations is the number of observations that can be obtained on a given sample in a realistic experimental time. To overcome this difficulty, we are planning to use the Mueller polarimetry to obtain at a fast rate (a few frames per second, compared to a few frames per half-hour) relevant information on the local anisotropy of biological (heart, skin) and composite (short fibers composite) samples.

G. Bureau, software engineer in the team, was funded by an Inria Reo industrial contract with Kephalios, a startup working on innovative artificial valves devices.

9.2. European Initiatives

9.2.1. Collaborations with Major European Organizations

Partner 1: Division of Biomedical Engineering & Imaging Sciences (BMEIS), St Thomas' Hospital, King's College London, UK

clinical-modeling topics mostly encompassing congenital heart diseases (BMEIS acts as "Other participant" in the Inria Associate team ToFMOD, and R. Chabiniok additionally performs clinical MRI exams at St Thomas' hospital 0.5 days / week.

Partner 2: Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Czech Republic

model-constrained image registrations, trans-valvular flow in pathological valves.

Partner 3: Institute for Clinical and Experimental Medicine in Prague

Cardiovascular MRI

9.3. International Initiatives

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

9.3.1.1. ToFMod

Title: Cardiac Biomechanical Modeling of Chronic Right Ventricular Loading

International Partner (Institution - Laboratory - Researcher):

UT Southwestern Medical Center, Dallas,Texas (United States), Mohammad Tarique Hussain

Start year: 2018

See also: https://m3disim.saclay.inria.fr/associated-team/

This collaboration aims at addressing a crucial issue in cardiology of congenital heart diseases, namely, the optimal timing of pulmonary valve replacement (PVR) in patients with surgically repaired tetralogy of Fallot (ToF) prone to chronic pulmonary regurgitation or right ventricular outflow tract stenosis. Our strategy consists in exploiting the predictive power of biomechanical modeling to shed light in the decision process. We will start by a detailed proof-of-concept study, based on datasets that will be acquired in patients indicated for percutaneous PVR, prior to the procedure, and in the follow-up at 3- and 12-months post-PVR. These datasets will be first used to calibrate the Inria M3DISIM patient-specific heart model simulating a cardiac cycle (at each follow-up time point) to access the myocardial properties – namely, the active contractility and passive stiffness. The instantaneous tissue properties will be statistically analyzed and compared with the level of reverse remodeling – i.e. the positive outcome of PVR. Secondly, the data at each time point will be used to calibrate and further develop the models of long-term tissue remodeling created by the M3DISIM researchers. It is only by combining such invaluable longitudinal data with biomechanical modeling expertise that progress can be achieved in the above objective, indeed.

9.4. International Research Visitors

9.4.1. Invited researchers

- T. Hussain, A. Tandon (Senior researchers at UTSW Medical Center Dallas): joint work in the scope of the Inria Associate team ToFMOD
- F. Regazzoni (3rd year PhD student from MOX, Milan, Italy): From Sept 2018, joint work on model learning and data assimilation coupling.

9.4.2. Internships

- K. Solovska (Czech Technical University and IKEM Prague): 1-30 August 2018, collaborative work with M. Genet and R. Chabiniok in the scope of the Inria Associate team ToFMOD

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific events organisation

10.1.1.1. General chair, scientific chair

R. Chabiniok, Organization of “MRI & Modelling workshop” related to the Inria Associate Team ToFMOD at Inria Paris-Saclay where hosting 14 international speakers from France, UK, USA, Switzerland, Chile, Czech Republic from the domains of MRI, cardiology and modelling (June 21)

R. Chabiniok, Co-chair of the section“Groupe de recherche en imagerie cardiaque” at French Congress of Radiology (Journées Francophones de Radiologie, JFR), Paris (October 12-15)

D. Chapelle, Co-chair of organising committee for VPH2020 Conference (Paris, August 2020)

D. Chapelle, Session co-chair in VPH18 Conference

M. Genet, Session co-chair at the 8th World Congress on Biomechanics, Dublin (July)

F. Clément, Session *Methodological developments for Systems Biology (A)*, ICSB 2018

P. Moireau, Session co-chair in VPH18 Conference

10.1.1.2. Member of organizing committees

M. Genet, Co-organizer of the Paris-Saclay University Biomechanics Seminar Series

A. Le Gall, Member of the organizing committee of WEARe 2018

A. Le Gall, Member of the organizing committee of the youth session of the national congress of the french society anesthesiology, intensive care, and peri-operative medicine

A. Le Gall, Member of the organizing committee of the youth days of teaching of the national congress of the french society anesthesiology, intensive care, and peri-operative medicine

P. Moireau, Member of the a organizing committee of the VPH2020 conference (Paris, August 2020)

P. Moireau, Member of the organizing committee of the Inria-Saclay teams (Poems-M3disim-Defi) scientific computing seminar

10.1.2. Scientific events selection

10.1.2.1. Reviewer

F. Clément *American Control Conference* 2019

P. Moireau, reviewer for VPH18

10.1.3. Journal

10.1.3.1. Member of editorial boards

D. Chapelle, Member of the editorial board of journal *Computers & Structures*

D. Chapelle, Member of the editorial board of journal *ESAIM: M2AN*

10.1.3.2. Reviewer - Reviewing activities

J.M. Allain, reviewer for “Acta Biomateriala”, “Journal of the Mechanical Behavior of Biomedical Materials” and “Journal of Applied Mathematics and Mechanics”

R. Chabiniok, reviewer for “Biomechanics and Modeling in Mechanobiology”, “Transations Of Society For Modeling And Simulations” and “Journal of Imaging”

D. Chapelle, reviewer for “Biomechanics and Modeling in Mechanobiology”, “Computers & Structures”, “International Journal for Numerical Methods in Biomedical Engineering”, “Meccanica”

F. Clément, Reviewer for “Journal of Mathematical Biology”, “PLOS Computational Biology”and “Endocrinology”

M. Genet, Reviewer for “Biomechanics and Modeling in Mechanobiology”

S. Imperiale, reviewer for “Journal of Mathematical Analysis”, “Journal of Computational Acoustics”, “SIAM Journal on Scientific Computing” and “SIAM Journal on Numerical Analysis”

P. Moireau, reviewer for “AMSES” and “IEEE Control Systems Letters”

F. Robin, Reviewer for “Journal of Mathematical Biology”

10.1.4. Invited talks

J. M. Allain, invited seminar at Century institut, France.

F. Caforio, seminar tours at the School of Biomedical Engineering and Imaging Sciences, King’s College London, UK, Simula Research Laboratory, Oslo, Norway and KFU university, Graz, Austria

R. Chabiniok, seminar tours at Institute for Clinical and Experimental Medicine, Mathematical Institute of Charles University and Children’s Heart Center at Motol University Hospital, Prague, Czech Republic

R. Chabiniok, invited lectures at Journées Francophones de Radiologie diagnostique et interventionnellle), France and at International Symposium on Modeling, Simulation and Optimization of the Cardiovascular System, Lukasklause Magdeburg, Germany

D. Chapelle, invited seminars at Politecnico di Milano (MOX), and for Dassault-Systèmes’ Living Heart Project

M. Genet, invited seminar at the School of Biomedical Engineering and Imaging Sciences, King’s College London, UK

S. Imperiale, invited lecture at ICERM, Brown University, USA

P. Le Tallec, seminar tours at University of Michigan, Michigan State University, Notre Dame University and University of Wisconsin at Madison, USA

P. Le Tallec, invited lecture at Charif University, Iran

P. Moireau, invited lectures at VPH18, CEMRACS2018 Luminy and FoMICS Summer School

P. Moireau, invited seminars at MOX, Politecnico di Milano, Italy and LMT, ENS Cachan, France

F. Robin invited seminar at Institut Élie Cartan de Lorraine, Nancy

10.1.5. Leadership within the scientific community

D. Chapelle, Member of the board of directors of the VPH Institute

D. Chapelle, Member of the steering committee of the BioMedical Engineering Institute coordinated by Ecole Polytechnique

F. Clément, member of the direction and scientific board of GdR REPRO

A. Le Gall, Chair of youth committee of SFAR (French Society of Anesthesia and Reanimation)

P. Le Tallec, Director of LMS (Solid Mechanics Laboratory) Ecole Polytechnique

P. Le Tallec, President of the Mechanics department at University Paris-Saclay

P. Moireau, Member of the steering committee of Department of Mathematics of Université Paris Saclay and Jacques Hadamard Foundation

10.1.6. Scientific expertise

R. Chabiniok, Honorary medical consultant at Saint-Thomas hospital (King’s College London)

R. Chabiniok, scientific consultant for the joint project of Institute for Clinical and Experimental Medicine in Prague (IKEM) and Czech Technical University

F. Clément, Reviewer for ANR

F. Clément, expert of the BCDE (Cell Biology, Development and Evolution) ITMO (Multi OrganizationThematic Institute) of the French National Alliance for Life and Health Sciences Aviesan.

F. Clément, expert for the INRA scientific prospective on predictive approaches in biology and ecology

S. Imperiale, Consultant for CEA

P. Le Tallec, Consultant for CEA

P. Le Tallec, Consultant for Michelin

P. Moireau, Reviewer for ANR

10.1.7. Research administration

J.M. Allain, Scientific Advisory Board, chair BioMecAM

R. Chabiniok, in charge of the objectives of Inria Associate team ToFMOD (with UT Southwestern Medical center Dallas, USA)

R. Chabiniok, in charge of coordination of clinical-modeling projects in the MΞDISIM team, Inria Saclay

D. Chapelle, Head of Science of Inria Saclay-Ile-de-France, and member of the Inria Evaluation Committee

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Bachelor: J.M. Allain, “Introductory Projects in Physics”, 10h, (L1), École Polytechnique, France

Bachelor: J.M. Allain, “Classical Mechanics”, 32h, (L2), École Polytechnique, France

Bachelor: F. Caforio,“PEIP1 S2 M2 – Mathematical analysis in two and three dimensions, linear algebra in Rn”, 22h, (L1), Polytech Paris-Sud, France

Bachelor: F. Caforio,“Math 255 – Differential calculus for physics (mathematical analysis in two and three dimensions)”, 42h, (L2), Université Paris-Sud, France

Bachelor: M. Genet, “Continuum Mechanics I”, 40h, (L3), École Polytechnique, France

Bachelor: M. Genet, “Modeling and Simulation in Industrial Mechanics”, 36h, (L3), École Polytechnique, France

Bachelor: S. Imperiale, “MA102 – Analyse pour les EDP”, 24h, (L3), ENSTA ParisTech, France

Bachelor: F. Kimmig, “Modeling and simulation in industrial mechanics”, 32h, (L3), École Polytechnique, France

Bachelor: P. Le Tallec, “MEC 431 – Mécanique des Milieux Continus 2”, (L3), École Polytechnique, France

Bachelor: P. Le Tallec, “Continuum mechanics”, 32h, (L3), Shanghai ParisTech, China

Bachelor: P. Moireau, “MAP 431 – Analyse variationnelle”, 40 h, (L3), Ecole Polytechnique, France

Bachelor: P. Moireau, “MODAL 472 – Expérimentation numérique pour les EDP”, 40 h, (L3), Ecole Polytechnique, France

Bachelor: F. Robin, “General mathematics”, 13h, (L1), Sorbonne Universités, France

Bachelor: F. Robin, “Probability and differentiel equation”, 22h (L2), Sorbonne Universités, France

Bachelor: F. Robin, “Power series and integral with parameters”, 20, (L2), Sorbonne Universités, France

Master: J.M. Allain, “Projects in Mechanics”, 6h, (M1), École Polytechnique, France

Master: J.M. Allain, “Cellular motility”, 32h, (M2), École Polytechnique, France

Master: J.M. Allain, “Supervision of the Experimental Center”, 45h, (M1), École Polytechnique, France

Master: D. Chapelle: “Biomechanical Modeling of Active Tissues”, 33h, (M2), Université Paris-Saclay, France

Master: M. Genet, “Plasticity and Fracture”, 18h, (M1), École Polytechnique, France

Master: S. Imperiale, “MA2610 Calcul Scientifique – Mécanique des solides”, 4h, (M1), Centrale/Supelec, France

Master: S. Imperiale, “Simnum – Programmation C++”, 18h, (M1), ENSTA ParisTech, France

Master: P. Le Tallec, “Nuclear Energy on Continuum Mechanics”,15h, (M2), INSTN, France

Master: P. Moireau, “Biomechanical Modeling of Active Tissues”, 12h, (M2), Université Paris-Saclay, France

Master: P. Moireau, “Méthodes et problèmes inverses en dynamique des populations”, 12h, (M2), UPMC, France

Master: P. Moireau, “AMS305 – Complétion de données et identification dans les problèmes gouvernés par des équations aux dérivées partielles”, 16h, (M2), Université Paris-Saclay, France

10.2.2. Supervision

PhD in progress: E. Berberoglu (ETHZ, Switzerland), “Image Guided Computational Cardiac Mechanics”, started February 2017, supervisors: M. Genet and S. Kozerke (ETHZ, Switzerland)

PhD in progress: F. Caforio, “Modélisation mathématique et numérique de la propagation d’ondes élastique dans le coeur”, started: Nov. 2015, supervisors: D. Chapelle and S. Imperiale

PhD in progress: C. Della Valle, “Modélisation et estimation des dynamiques d’assemblage de proteines”, supervisors: M. Doumic and P. Moireau, Université Paris Sciences et Lettres,

PhD in progress: C. Giraudet, “Cornea biomechanics”, started 10/2018; supervisors: J.M. Allain and P. Le Tallec

PhD in progress: M. Gusseva, “Cardiac Biomechanical Modeling for Chronic Ventricular Loading”, supervisors: R. Chabiniok, D. Chapelle, T. Hussain, Université Paris-Saclay, started in December 2017

PhD: U. Khristenko, Université Paris-Saclay, P. Le Tallec, defended Jan 17th

PhD in progress: F. Kimmig, “Multi-scale modeling of muscle contraction – From stochastic dynamics of molecular motors to continuum mechanics, in interaction with experimental assays”, supervisors: M. Caruel and D. Chapelle, Université Paris-Saclay, started in September 2016

PhD in progress: A. Le Gall, “Cardiac modelling for monitoring purposes during general anaesthesia and at Intensive Care Unit”, supervisors: R. Chabiniok, D. Chapelle, E. Gayat, Université Paris-Saclay, started in November 2016

PhD in progress: C. Patte, “Lung multiscale poromechanical modeling: from breathing to pulmonary fibrosis-induced chronic remodeling”, started November 2017, supervisors: M. Genet and D. Chapelle

PhD in progress: F. Robin “Multiscale modeling of the morphodynamics in ovarian follicles”, started October 2016, supervisors: F. Clément & Romain Yvinec [INRA]

PhD in progress: N. Tueni, “Multiscale modeling of cardiac mechanics”, started January 2018, supervisors: M. Genet and J.-M. Allain

PhD in progress: F. Wijanto, “Modélisation multi-échelle des fibres de collagènes”, started: Sept. 2015, supervisors: J.-M. Allain and M. Caruel

Supervision of a project research in CEMRACS 2018 “ Multiscale population dynamics : interactions between scales in developmental and reproductive biology” (F. Clément, M. Postel and R. Yvinec)

10.2.3. Juries

D. Chapelle, PhD Jury of A. Marboeuf, Ecole Polytechnique, March 8

P. Le Tallec, PhD reviewer of R. Mlika, INSA Lyon, Jan 24

P. Le Tallec, PhD Jury of Q. Pierron, ENSTA ParisTech, May 18

P. Le Tallec, PhD reviewer of L. Poirel, Inria Bordeaux, Nov 28

P. Moireau, PhD Jury of U. Khristenko, Université Paris-Saclay, Jan 17th

P. Moireau, PhD Jury of T. Kritter, Bordeaux University, Oct 1st

10.3. Popularization

J.M. Allain, Co-authored a popularization paper in the “Reflets de la Physique” journal (january-february 2018)

D. Chapelle, Interview in *Sciences et Avenir* (November issue)

M. Genet, Presentation at JeudiX—Research Days at École Polytechnique

C. Patte, Fête des Sciences day, organized by Inria. October 11th

11. Bibliography

Major publications by the team in recent years

[1] J. ALBELLA MARTÍNEZ, S. IMPERIALE, P. JOLY, J. RODRÍGUEZ.*Solving 2D linear isotropic elastodynamics by means of scalar potentials: a new challenge for finite elements*, in "Journal of Scientific Computing", 2018 [*DOI :* 10.1007/s10915-018-0768-9], https://hal.inria.fr/hal-01803536

[2] M. CARUEL, P. MOIREAU, D. CHAPELLE.*Stochastic modeling of chemical-mechanical coupling in striated muscles*, in "Biomechanics and Modeling in Mechanobiology", 2018 [*DOI :* 10.1007/s10237-018-1102-z], https://hal.inria.fr/hal-01928279

[3] R. CHABINIOK, P. MOIREAU, P.-F. LESAULT, A. RAHMOUNI, J.-F. DEUX, D. CHAPELLE.*Estimation of tissue contractility from cardiac cine-MRI using a biomechanical heart model*, in "Biomechanics and Modeling in Mechanobiology", 2012, vol. 11, n^{o} 5, p. 609-630 [*DOI :* 10.1007/s10237-011-0337-8], http://hal.inria.fr/hal-00654541

[4] D. CHAPELLE, K. BATHE.*The Finite Element Analysis of Shells - Fundamentals - Second Edition*, Computational Fluid and Solid Mechanics, Springer, 2011, 410 [*DOI :* 10.1007/978-3-642-16408-8], http://hal.inria.fr/hal-00654533

[5] D. CHAPELLE, N. CÎNDEA, P. MOIREAU.*Improving convergence in numerical analysis using observers - The wave-like equation case*, in "Mathematical Models and Methods in Applied Sciences", 2012, vol. 22, n^{o} 12 [*DOI :* 10.1142/S0218202512500406], http://hal.inria.fr/inria-00621052

[6] D. CHAPELLE, P. LE TALLEC, P. MOIREAU, M. SORINE.*An energy-preserving muscle tissue model: formulation and compatible discretizations*, in "International Journal for Multiscale Computational Engineering", 2012, vol. 10, n^{o} 2, p. 189-211 [*DOI :* 10.1615/INTJMULTCOMPENG.2011002360], http://hal.inria.fr/hal-00678772

[7] D. CHAPELLE, P. MOIREAU.*General coupling of porous flows and hyperelastic formulations – From thermodynamics principles to energy balance and compatible time schemes*, in "European Journal of Mechanics - B/Fluids", 2014, vol. 46, p. 82-96, Updated version of previously published research report [*DOI :* 10.1016/J.EUROMECHFLU.2014.02.009], https://hal.inria.fr/inria-00520612

[8] A. COLLIN, S. IMPERIALE.*Mathematical analysis and 2-scale convergence of a heterogeneous microscopic bidomain model*, in "Mathematical Models and Methods in Applied Sciences", 2018, https://hal.inria.fr/hal-01759914

[9] B. LYNCH, S. BANCELIN, C. BONOD-BIDAUD, J.-B. GUEUSQUIN, F. RUGGIERO, M.-C. SCHANNE-KLEIN, J.-M. ALLAIN.*A novel microstructural interpretation for the biomechanics of mouse skin derived from multiscale characterization*, in "Acta Biomaterialia", 2017, vol. 50, p. 302-311 [*DOI :* 10.1016/J.ACTBIO.2016.12.051], https://hal.archives-ouvertes.fr/hal-01531321

[10] P. MOIREAU.*A Discrete-time Optimal Filtering Approach for Non-linear Systems as a Stable Discretization of the Mortensen Observer*, in "ESAIM: Control, Optimisation and Calculus of Variations", 2017, https://hal.inria.fr/hal-01671271

[11] M. SERMESANT, R. CHABINIOK, P. CHINCHAPATNAM, T. MANSI, F. BILLET, P. MOIREAU, J.-M. PEYRAT, K. C. WONG, J. RELAN, K. S. RHODE, M. GINKS, P. LAMBIASE, H. DELINGETTE, M. SORINE, C. A. RINALDI, D. CHAPELLE, R. RAZAVI, N. AYACHE.*Patient-Specific Electromechanical Models of the Heart for Prediction of the Acute Effects of Pacing in CRT: a First Validation*, in "Medical Image Analysis", 2012, vol. 16, n^{o} 1, p. 201-215 [*DOI :* 10.1016/J.MEDIA.2011.07.003], http://hal.inria.fr/inria-00616191

Publications of the year

Articles in International Peer-Reviewed Journal

[12] A. AALTO.*Convergence of discrete-time Kalman filter estimate to continuous-time estimate for systems with unbounded observation*, in "Mathematics of Control, Signals, and Systems", June 2018, vol. 30, n^{o} 3, 9, https://arxiv.org/abs/1512.02473 - This is the preprint version of the article published in Mathematics of Control, Signals and Systems [*DOI :* 10.1007/S0049], https://hal.inria.fr/hal-01236950

[13] A. AALTO.*Iterative observer-based state and parameter estimation for linear systems*, in "ESAIM: Control, Optimisation and Calculus of Variations", May 2018, vol. 24, n^{o} 1, p. 265-288, The original publication is available at www.esaim-cocv.org [*DOI :* 10.1051/COCV/2017005], https://hal.inria.fr/hal-01370430

[14] J. ALBELLA MARTÍNEZ, S. IMPERIALE, P. JOLY, J. RODRÍGUEZ.*Solving 2D linear isotropic elastodynamics by means of scalar potentials: a new challenge for finite elements*, in "Journal of Scientific Computing", 2018 [*DOI :* 10.1007/S10915-018-0768-9], https://hal.inria.fr/hal-01803536

[15] A. BOUGLÉ, A. LE GALL, F. DUMAS, G. GERI, I. MALISSIN, S. VOICU, B. MÉGARBANE, A. CARIOU, N. DEYE.*Extra Corporeal Life Support for Cardiac ARrest patients with post-cardiac arrest syndrome: the ECCAR study*, in "Archives of cardiovascular diseases", 2018, https://hal.archives-ouvertes.fr/hal-01942222

[16] B. Burtschell, P. Moireau, D. Chapelle.*Numerical analysis for an energy-stable total discretization of a poromechanics model with inf-sup stability*, in "Acta Mathematicae Applicatae Sinica", 2018, https://hal.inria.fr/hal-01928312

[17] F. Caforio, S. Imperiale.*A conservative penalisation strategy for the semi-implicit time discretisation of incompressible elastodynamics equation*, in "Advanced Modeling and Simulation in Engineering Sciences", 2018, https://hal.inria.fr/hal-01907746

[18] M. Caruel, P. Moireau, D. Chapelle.*Stochastic modeling of chemical-mechanical coupling in striated muscles*, in "Biomechanics and Modeling in Mechanobiology", 2018 [*DOI :* 10.1007/s10237-018-1102-z], https://hal.inria.fr/hal-01928279

[19] A. Collin, S. Imperiale.*Mathematical analysis and 2-scale convergence of a heterogeneous microscopic bidomain model*, in "Mathematical Models and Methods in Applied Sciences", 2018, https://hal.inria.fr/hal-01759914

[20] J.-L. FELLAHI, A. Godier, D. Benchetrit, F. Berthier, G. Besch, T. Bochaton, E. Bonnefoy-Cudraz, P. Coriat, E. Gayat, A. Hong, S. Jenck, A. Le Gall, D. Longrois, A.-C. Martin, S. Pili-Floury, V. Piriou, S. Provenchère, B. Rozec, E. Samain, R. SCHWEIZER, V. Billard.*Perioperative management of patients with coronary artery disease undergoing non-cardiac surgery: Summary from the French Society of Anaesthesia and Intensive Care Medicine 2017 convention*, in "Anaesthesia Critical Care & Pain Medicine", August 2018, vol. 37, n^o 4, p. 367 - 374 [*DOI :* 10.1016/j.accpm.2018.02.021], https://hal.archives-ouvertes.fr/hal-01927277

[21] H. Finsberg, C. Xi, J. L. Tan, L. Zhong, M. Genet, J. Sundnes, L. C. Lee, S. Wall.*Efficient estimation of personalized biventricular mechanical function employing gradient-based optimization*, in "International Journal for Numerical Methods in Biomedical Engineering", 2018, vol. 34, n° 7, e2982 [*DOI :* 10.1002/cnm.2982], https://hal.archives-ouvertes.fr/hal-01882363

[22] M. Genet, C. T. Stoeck, C. von Deuster, L. C. Lee, S. Kozerke.*Equilibrated warping: Finite element image registration with finite strain equilibrium gap regularization*, in "Medical Image Analysis", 2018, vol. 50, p. 1-22 [*DOI :* 10.1016/j.media.2018.07.007], https://hal.archives-ouvertes.fr/hal-01882366

[23] J. Hoermann, C. Bertoglio, M. Kronbichler, M. Pfaller, R. Chabiniok, W. Wall.*An adaptive Hybridizable Discontinuous Galerkin approach for cardiac electrophysiology*, in "International Journal for Numerical Methods in Biomedical Engineering", 2018, vol. 34, n° 5, e2959 (20 pages) [*DOI :* 10.1002/cnm.2959], https://hal.archives-ouvertes.fr/hal-01672964

[24] U. Khristenko, P. Le Tallec.*Delayed feedback control method for computing the cyclic steady states of evolution problems*, in "Computer Methods in Applied Mechanics and Engineering", August 2018, vol. 338, p. 241 - 274 [*DOI :* 10.1016/j.cma.2018.04.015], https://hal.archives-ouvertes.fr/hal-01819125

[25] U. Khristenko, P. Le Tallec.*Newton-Krylov method for computing the cyclic steady states of evolution problems in non-linear mechanics*, in "International Journal for Numerical Methods in Engineering", 2018 [*DOI :* 10.1002/nme.5920], https://hal.archives-ouvertes.fr/hal-01830166

[26] M. Pfaller, J. Hoermann, M. Weigl, A. Nagler, R. Chabiniok, C. Bertoglio, W. A. Wall.*The importance of the pericardium for cardiac biomechanics: From physiology to computational modeling*, in

"Biomechanics and Modeling in Mechanobiology", 2018, https://arxiv.org/abs/1810.05451 , https://hal.inria.fr/hal-01941544

[27] P. SHIRKOVSKIY, A. LAURIN, N. JEGER-MADIOT, D. CHAPELLE, M. FINK, R. K. ING.*Airborne ultrasound surface motion camera: application to seismocardiography*, in "Applied Physics Letters", 2018, vol. 112 [*DOI :* 10.1063/1.5028348], https://hal.inria.fr/hal-01791827

[28] A. TANTOT, A. CAILLARD, A. LE GALL, J. MATÉO, S. MILLASSEAU, A. MEBAZAA, E. GAYAT, F. VALLÉE.*Evaluation of cardiac output variations with the peripheral pulse pressure to mean arterial pressure ratio*, in "Journal of Clinical Monitoring and Computing", October 2018 [*DOI :* 10.1007/S10877-018-0210-8], https://hal.archives-ouvertes.fr/hal-01927292

[29] J. WONG, R. CHABINIOK, S. M. TIBBY, K. PUSHPARAJAH, E. SAMMUT, D. CELERMAJER, D. GIESE, T. HUSSAIN, G. F. GREIL, T. SCHAEFFTER, R. RAZAVI.*Exploring kinetic energy as a new marker of cardiac function in the single ventricle circulation*, in "Journal of Applied Physiology", January 2018, vol. 125, p. 889-900 [*DOI :* 10.1152/JAPPLPHYSIOL.00580.2017], https://hal.archives-ouvertes.fr/hal-01697067

[30] R. YVINEC, P. CREPIEUX, E. REITER, A. POUPON, F. CLÉMENT.*Advances in computational modeling approaches of pituitary gonadotropin signaling*, in "Expert Opinion on Drug Discovery", 2018, vol. 13, n^{o} 9, p. 799-813 [*DOI :* 10.1080/17460441.2018.1501025], https://hal.archives-ouvertes.fr/hal-01913612

[31] C. ZHANG, A. COLLIN, P. MOIREAU, A. TROUVÉ, M. ROCHOUX.*Front shape similarity measure for data-driven simulations of wildland fire spread based on state estimation: Application to the RxCADRE field-scale experiment*, in "Proceedings of the Combustion Institute", September 2018, https://hal.inria.fr/hal-01969130

[32] H. ZOU, C. XI, X. ZHAO, A. S. KOH, F. GAO, Y. SU, R.-S. Y. TAN, J. ALLEN, L. C. LEE, M. GENET, L. ZHONG.*Quantification of Biventricular Strains in Heart Failure With Preserved Ejection Fraction Patient Using Hyperelastic Warping Method*, in "Frontiers in Physiology", September 2018, vol. 9 [*DOI :* 10.3389/FPHYS.2018.01295], https://hal.archives-ouvertes.fr/hal-01882365

Conferences without Proceedings

[33] M. GENET.*A continuum relaxed growth model for controlling growth-induced residual stresses in living tissues*, in "8th World Congress of Biomechanics", Dublin, Ireland, July 2018, https://hal.archives-ouvertes.fr/hal-01882372

[34] M. GENET, C. T. STOECK, E. BERBEROGLU, S. KOZERKE.*Equilibrated Warping: Finite Element Image Registration with Equilibrium Gap Regularization*, in "8th World Congress of Biomechanics", Dublin, Ireland, July 2018, https://hal.archives-ouvertes.fr/hal-01882370

Research Reports

[35] H. NGUYEN, F. ROBIN, A. VONSEEL.*Transformation paramétrique de distance à vol d'oiseau en distance sur une voirie*, AMIES, December 2018, Semaine d'Etude Maths-Entreprises,Poitiers, du 28 au 1er Juin 2018, https://hal.archives-ouvertes.fr/hal-01949328

Other Publications

[36] J. ALBELLA MARTÍNEZ, H. BEN DHIA, S. IMPERIALE, J. RODRÍGUEZ.*Mathematical and numerical study of transient wave scattering by obstacles with the Arlequin Method*, October 2018, working paper or preprint, https://hal.inria.fr/hal-01898420

[37] D. H. BAFFET, M. J. GROTE, S. IMPERIALE, M. KACHANOVSKA.*Energy Decay and Stability of a Perfectly Matched Layer For the Wave Equation*, September 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01865484

[38] F. CAFORIO, S. IMPERIALE.*High-order discrete fourier transform for the solution of the Poisson equation*, November 2018, working paper or preprint, https://hal.inria.fr/hal-01914257

[39] J. CHABASSIER, J. DIAZ, S. IMPERIALE.*Construction and analysis of fourth order, energy consistent, family of explicit time discretizations for dissipative linear wave equations*, October 2018, working paper or preprint, https://hal.inria.fr/hal-01894238

[40] J. CHABASSIER, S. IMPERIALE.*Construction and convergence analysis of conservative second order local time discretisation for wave equations*, October 2018, working paper or preprint, https://hal.inria.fr/hal-01894357

[41] F. CLÉMENT, F. ROBIN, R. YVINEC.*Analysis and calibration of a linear model for structured cell populations with unidirectional motion : Application to the morphogenesis of ovarian follicles*, August 2018, https://arxiv.org/abs/1712.05372 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01852560

[42] S. IMPERIALE, P. MOIREAU, A. TONNOIR.*Analysis of an observer strategy for initial state reconstruction of wave-like systems in unbounded domains*, October 2018, working paper or preprint, https://hal.inria.fr/hal-01898440

[43] M. POSTEL, A. KARAM, G. PÉZERON, S. SCHNEIDER-MAUNOURY, F. CLÉMENT.*A multiscale mathematical model of cell dynamics during neurogenesis in the mouse cerebral cortex*, May 2018, working paper or preprint, https://hal.inria.fr/hal-01783141

Project-Team MEXICO

Modeling and Exploitation of Interaction and Concurrency

IN COLLABORATION WITH: Laboratoire specification et vérification (LSV)

IN PARTNERSHIP WITH:
CNRS

Ecole normale supérieure de Cachan

RESEARCH CENTER
Saclay - Île-de-France

THEME
Proofs and Verification

Table of contents

Project-Team MEXICO

Creation of the Team: 2009 March 01, updated into Project-Team: 2011 January 01

Keywords:

Computer Science and Digital Science:

A2.3. - Embedded and cyber-physical systems
A2.3.2. - Cyber-physical systems
A2.3.3. - Real-time systems
A2.4.1. - Analysis
A2.4.2. - Model-checking
A6.4.1. - Deterministic control
A6.4.3. - Observability and Controlability
A7.1. - Algorithms
A7.1.1. - Distributed algorithms
A7.2. - Logic in Computer Science
A7.3.1. - Computational models and calculability
A8.1. - Discrete mathematics, combinatorics
A8.2. - Optimization
A8.7. - Graph theory
A8.8. - Network science
A8.9. - Performance evaluation
A8.11. - Game Theory

Other Research Topics and Application Domains:

B1.1.2. - Molecular and cellular biology
B1.1.10. - Systems and synthetic biology
B6.3.1. - Web
B6.3.2. - Network protocols
B6.3.3. - Network Management
B7.1. - Traffic management
B7.2.1. - Smart vehicles

1. Team, Visitors, External Collaborators

Research Scientists

Stefan Haar [Team leader, Inria, Senior Researcher, HDR]
Laurent Fribourg [CNRS, Researcher, HDR]
Matthias Függer [CNRS, Researcher]

Faculty Members

Thomas Chatain [Ecole Normale Supérieure Cachan, Associate Professor]
Philippe Dague [Univ Paris-Sud, Professor, from Oct 2018]
Serge Haddad [Ecole Normale Supérieure Cachan, Professor, HDR]
Claudine Picaronny [Ecole Normale Supérieure Cachan, Professor, until Jul 2018]
Stefan Schwoon [Ecole Normale Supérieure Cachan, Associate Professor]

External Collaborator
Benoît Barbot [Univ Paris-Est Marne La Vallée]

PhD Students
Mathilde Boltenhagen [CNRS, from Nov 2018]
Yann Duplouy [Institut de recherche technologique System X, until Nov 2018]
Jawher Jerray [Univ Paris-Nord, from Oct 2018]
Igor Khmelnitsky [Ecole Normale Supérieure Cachan, from Nov 2018]
Juraj Kolcˇák [Inria]
Hugues Mandon [Inria]
Adnane Saoud [CNRS]

Post-Doctoral Fellows
Da Jung Cho [CNRS, from May 2018]
Daniele Zonetti [CNRS, from May 2018]

Administrative Assistants
Adeline Lochet [Inria, from Jun 2018 until Nov 2018]
Emmanuelle Perrot [Inria]

2. Overall Objectives

2.1. Scientific Objectives

2.1.1. Introduction.

In the increasingly networked world, reliability of applications becomes ever more critical as the number of users of, e.g., communication systems, web services, transportation etc., grows steadily. Management of networked systems, in a very general sense of the term, therefore is a crucial task, but also a difficult one.

MExICo strives to take advantage of distribution by orchestrating cooperation between different agents that observe local subsystems, and interact in a localized fashion.

The need for applying formal methods in the analysis and management of complex systems has long been recognized. It is with much less unanimity that the scientific community embraces methods based on asynchronous and distributed models. Centralized and sequential modeling still prevails.

However, we observe that crucial applications have increasing numbers of users, that networks providing services grow fast both in the number of participants and the physical size and degree of spatial distribution. Moreover, traditional *isolated* and *proprietary* software products for local systems are no longer typical for emerging applications.

In contrast to traditional centralized and sequential machinery for which purely functional specifications are efficient, we have to account for applications being provided from diverse and non-coordinated sources. Their distribution (e.g. over the Web) must change the way we verify and manage them. In particular, one cannot ignore the impact of quantitative features such as delays or failure likelihoods on the functionalities of composite services in distributed systems.

We thus identify three main characteristics of complex distributed systems that constitute research challenges:

- *Concurrency* of behavior;
- *Interaction* of diverse and semi-transparent components; and
- management of *Quantitative* aspects of behavior.

2.1.2. Concurrency

The increasing size and the networked nature of communication systems, controls, distributed services, etc. confront us with an ever higher degree of parallelism between local processes. This field of application for our work includes telecommunication systems and composite web services. The challenge is to provide sound theoretical foundations and efficient algorithms for management of such systems, ranging from controller synthesis and fault diagnosis to integration and adaptation. While these tasks have received considerable attention in the *sequential* setting, managing *non-sequential* behavior requires profound modifications for existing approaches, and often the development of new approaches altogether. We see concurrency in distributed systems as an opportunity rather than a nuisance. Our goal is to *exploit* asynchronicity and distribution as an advantage. Clever use of adequate models, in particular *partial order semantics* (ranging from Mazurkiewicz traces to event structures to MSCs) actually helps in practice. In fact, the partial order vision allows us to make causal precedence relations explicit, and to perform diagnosis and test for the dependency between events. This is a conceptual advantage that interleaving-based approaches cannot match. The two key features of our work will be *(i)* the exploitation of concurrency by using asynchronous models with partial order semantics, and *(ii)* distribution of the agents performing management tasks.

2.1.3. Interaction

Systems and services exhibit non-trivial *interaction* between specialized and heterogeneous components. A coordinated interplay of several components is required; this is challenging since each of them has only a limited, partial view of the system's configuration. We refer to this problem as *distributed synthesis* or *distributed control*. An aggravating factor is that the structure of a component might be semi-transparent, which requires a form of *grey box management*.

2.1.4. Quantitative Features

Besides the logical functionalities of programs, the *quantitative* aspects of component behavior and interaction play an increasingly important role.

- *Real-time* properties cannot be neglected even if time is not an explicit functional issue, since transmission delays, parallelism, etc, can lead to time-outs striking, and thus change even the logical course of processes. Again, this phenomenon arises in telecommunications and web services, but also in transport systems.
- In the same contexts, *probabilities* need to be taken into account, for many diverse reasons such as unpredictable functionalities, or because the outcome of a computation may be governed by race conditions.
- Last but not least, constraints on *cost* cannot be ignored, be it in terms of money or any other limited resource, such as memory space or available CPU time.

2.1.5. Evolution and Perspectives

Since the creation of *MExICo*, the weight of *quantitative* aspects in all parts of our activities has grown, be it in terms of the models considered (weighted automata and logics), be it in transforming verification or diagnosis verdict into probabilistic statements (probabilistic diagnosis, statistical model checking), or within the recently started SystemX cooperation on supervision in multi-modal transport systems. This trend is certain to continue over the next couple of years, along with the growing importance of diagnosis and control issues.

In another development, the theory and use of partial order semantics has gained momentum in the past four years, and we intend to further strengthen our efforts and contacts in this domain to further develop and apply partial-order based deduction methods.

When no complete model of the underlying dynamic system is available, the analysis of logs may allow to reconstruct such a model, or at least to infer some properties of interest; this activity, which has emerged over the past 10 years on the international level, is referred to as **process mining**. In this emerging activity, we have contributed to unfolding-based process discovery [CI-146], and the study of process alignments [CI-121, CI-96, CI-83, CI-60, CI-33].

Finally, over the past years *biological* challenges have come to the center of our work, in two different directions:

1. **(Re-)programming in discrete concurrent models.** Cellular regulatory networks exhibit highly complex concurrent behaviours that is influenced by a high number of perturbations such as mutations. We are in particular investigating discrete models, both in the form of boolean networks and of Petri nets, to harness this complexity, and to obtain viable methods for two interconnected and central challenges:
   - find *attractors*, i.e. long-run stable states or sets of states, that indicate possible phenotypes of the organism under study, and
   - determine *reprogramming* strategies that apply perturbations in such a way as to steer the cell's long-run behaviour into some desired phenotype, or away from an undesired one.
2. **Distributed Algorithms in wild or synthetic biological systems.** Since the arrival of Matthias Függer in the team, we also work, on the multi-cell level, with a distributed algorithms' view on microbiological systems, both with the goal to model and analyze existing microbiological systems as distributed systems, and to design and implement distributed algorithms in synthesized microbiological systems. Major long-term goals are drug production and medical treatment via synthesized bacterial colonies.

3. Research Program

3.1. Concurrency

Participants: Thomas Chatain, Stefan Haar, Serge Haddad, Stefan Schwoon.

Concurrency: Property of systems allowing some interacting processes to be executed in parallel.

Diagnosis: The process of deducing from a partial observation of a system aspects of the internal states or events of that system; in particular, *fault diagnosis* aims at determining whether or not some non-observable fault event has occurred.

Conformance Testing: Feeding dedicated input into an implemented system IS and deducing, from the resulting output of I, whether I respects a formal specification S.

3.1.1. Introduction

It is well known that, whatever the intended form of analysis or control, a *global* view of the system state leads to overwhelming numbers of states and transitions, thus slowing down algorithms that need to explore the state space. Worse yet, it often blurs the mechanics that are at work rather than exhibiting them. Conversely, respecting concurrency relations avoids exhaustive enumeration of interleavings. It allows us to focus on 'essential' properties of non-sequential processes, which are expressible with causal precedence relations. These precedence relations are usually called causal (partial) orders. Concurrency is the explicit absence of such a precedence between actions that do not have to wait for one another. Both causal orders and concurrency are in fact essential elements of a specification. This is especially true when the specification is constructed in a distributed and modular way. Making these ordering relations explicit requires to leave the framework of state/interleaving based semantics. Therefore, we need to develop new dedicated algorithms for tasks such as conformance testing, fault diagnosis, or control for distributed discrete systems. Existing solutions for these problems often rely on centralized sequential models which do not scale up well.

3.1.2. Diagnosis

Participants: Stefan Haar, Serge Haddad, Stefan Schwoon.

Fault Diagnosis for discrete event systems is a crucial task in automatic control. Our focus is on *event oriented* (as opposed to *state oriented*) model-based diagnosis, asking e.g. the following questions:

given a - potentially large - *alarm pattern* formed of observations,

- what are the possible *fault scenarios* in the system that *explain* the pattern ?
- Based on the observations, can we deduce whether or not a certain - invisible - fault has actually occurred ?

Model-based diagnosis starts from a discrete event model of the observed system - or rather, its relevant aspects, such as possible fault propagations, abstracting away other dimensions. From this model, an extraction or unfolding process, guided by the observation, produces recursively the explanation candidates.

In asynchronous partial-order based diagnosis with Petri nets [49], [50], [51], one unfolds the *labelled product* of a Petri net model $\mathcal{N}$ and an observed alarm pattern $\mathcal{A}$, also in Petri net form. We obtain an acyclic net giving partial order representation of the behaviors compatible with the alarm pattern. A recursive online procedure filters out those runs *(configurations)* that explain *exactly*$\mathcal{A}$. The Petri-net based approach generalizes to dynamically evolving topologies, in dynamical systems modeled by graph grammars, see [38]

3.1.2.1. Observability and Diagnosability

Diagnosis algorithms have to operate in contexts with low observability, i.e., in systems where many events are invisible to the supervisor. Checking *observability* and *diagnosability* for the supervised systems is therefore a crucial and non-trivial task in its own right. Analysis of the relational structure of occurrence nets allows us to check whether the system exhibits sufficient visibility to allow diagnosis. Developing efficient methods for both verification of *diagnosability checking* under concurrency, and the *diagnosis* itself for distributed, composite and asynchronous systems, is an important field for *MExICo*.

3.1.2.2. Distribution

Distributed computation of unfoldings allows one to factor the unfolding of the global system into smaller *local* unfoldings, by local supervisors associated with sub-networks and communicating among each other. In [50], [40], elements of a methodology for distributed computation of unfoldings between several supervisors, underwritten by algebraic properties of the category of Petri nets have been developed. Generalizations, in particular to Graph Grammars, are still do be done.

Computing diagnosis in a distributed way is only one aspect of a much vaster topic, that of *distributed diagnosis* (see [47], [53]). In fact, it involves a more abstract and often indirect reasoning to conclude whether or not some given invisible fault has occurred. Combination of local scenarios is in general not sufficient: the global system may have behaviors that do not reveal themselves as faulty (or, dually, non-faulty) on any local supervisor's domain (compare [37], [43]). Rather, the local diagnosers have to join all *information* that is available to them locally, and then deduce collectively further information from the combination of their views. In particular, even the *absence* of fault evidence on all peers may allow to deduce fault occurrence jointly, see [55], [56]. Automatizing such procedures for the supervision and management of distributed and locally monitored asynchronous systems is a long-term goal to which *MExICo* hopes to contribute.

3.1.3. Hybrid Systems

Participants: Laurent Fribourg, Serge Haddad.

Hybrid systems constitute a model for cyber-physical systems which integrates continuous-time dynamics (modes) governed by differential equations, and discrete transitions which switch instantaneously from one mode to another. Thanks to their ease of programming, hybrid systems have been integrated to power electronics systems, and more generally in cyber-physical systems. In order to guarantee that such systems meet their specifications, classical methods consist in finitely abstracting the systems by discretization of the (infinite) state space, and deriving automatically the appropriate mode control from the specification using standard graph techniques. These methods face the well-known problem of “curse of dimensionality”, and cannot generally treat systems of dimension exceeding 5 or 6. Thanks to the introduction of original compositional techniques [25], [30], [13] as well as finer estimations of integration errors [3], we are now able to control several case studies of greater dimension. Actually, in the real world, many parameters of hybrid models are not known precisely, and require adjustements to experimental data. We plan to elaborate methods based on parameter estimation and machine learning techniques in order to define formal stability criteria and well-posed learning problems in the framework of hybrid systems with nonlinear dynamics.

3.1.4. Contextual Nets

Participant: Stefan Schwoon.

Assuring the correctness of concurrent systems is notoriously difficult due to the many unforeseeable ways in which the components may interact and the resulting state-space explosion. A well-established approach to alleviate this problem is to model concurrent systems as Petri nets and analyse their unfoldings, essentially an acyclic version of the Petri net whose simpler structure permits easier analysis [48].

However, Petri nets are inadequate to model concurrent read accesses to the same resource. Such situations often arise naturally, for instance in concurrent databases or in asynchronous circuits. The encoding tricks typically used to model these cases in Petri nets make the unfolding technique inefficient. Contextual nets, which explicitly do model concurrent read accesses, address this problem. Their accurate representation of concurrency makes contextual unfoldings up to exponentially smaller in certain situations. An abstract algorithm for contextual unfoldings was first given in [39]. In recent work, we further studied this subject from a theoretical and practical perspective, allowing us to develop concrete, efficient data structures and algorithms and a tool (Cunf) that improves upon existing state of the art. This work led to the PhD thesis of César Rodríguez in 2014 .

Contextual unfoldings deal well with two sources of state-space explosion: concurrency and shared resources. Recently, we proposed an improved data structure, called *contextual merged processes* (CMP) to deal with a third source of state-space explosion, i.e. sequences of choices. The work on CMP [57] is currently at an abstract level. In the short term, we want to put this work into practice, requiring some theoretical groundwork, as well as programming and experimentation.

Another well-known approach to verifying concurrent systems is *partial-order reduction*, exemplified by the tool SPIN. Although it is known that both partial-order reduction and unfoldings have their respective strengths and weaknesses, we are not aware of any conclusive comparison between the two techniques. Spin comes with a high-level modeling language having an explicit notion of processes, communication channels, and variables. Indeed, the reduction techniques implemented in Spin exploit the specific properties of these features. On the other side, while there exist highly efficient tools for unfoldings, Petri nets are a relatively general low-level formalism, so these techniques do not exploit properties of higher language features. Our work on contextual unfoldings and CMPs represents a first step to make unfoldings exploit richer models. In the long run, we wish raise the unfolding technique to a suitable high-level modelling language and develop appropriate tool support.

3.2. Management of Quantitative Behavior

Participants: Thomas Chatain, Stefan Haar, Serge Haddad.

3.2.1. Introduction

Besides the logical functionalities of programs, the *quantitative* aspects of component behavior and interaction play an increasingly important role.

- *Real-time* properties cannot be neglected even if time is not an explicit functional issue, since transmission delays, parallelism, etc, can lead to time-outs striking, and thus change even the logical course of processes. Again, this phenomenon arises in telecommunications and web services, but also in transport systems.
- In the same contexts, *probabilities* need to be taken into account, for many diverse reasons such as unpredictable functionalities, or because the outcome of a computation may be governed by race conditions.
- Last but not least, constraints on *cost* cannot be ignored, be it in terms of money or any other limited resource, such as memory space or available CPU time.

Traditional mainframe systems were proprietary and (essentially) localized; therefore, impact of delays, unforeseen failures, etc. could be considered under the control of the system manager. It was therefore natural, in verification and control of systems, to focus on *functional* behavior entirely.

With the increase in size of computing system and the growing degree of compositionality and distribution, quantitative factors enter the stage:

- calling remote services and transmitting data over the web creates *delays*;
- remote or non-proprietary components are not "deterministic", in the sense that their behavior is uncertain.

Time and *probability* are thus parameters that management of distributed systems must be able to handle; along with both, the *cost* of operations is often subject to restrictions, or its minimization is at least desired. The mathematical treatment of these features in distributed systems is an important challenge, which *MExICo* is addressing; the following describes our activities concerning probabilistic and timed systems. Note that cost optimization is not a current activity but enters the picture in several intended activities.

3.2.2. Probabilistic distributed Systems

Participants: Stefan Haar, Serge Haddad.

3.2.2.1. Non-sequential probabilistic processes

Practical fault diagnosis requires to select explanations of *maximal likelihood*. For partial-order based diagnosis, this leads therefore to the question what the probability of a given partially ordered execution is. In Benveniste et al. [42], [35], we presented a model of stochastic processes, whose trajectories are partially ordered, based on local branching in Petri net unfoldings; an alternative and complementary model based on Markov fields is developed in [52], which takes a different view on the semantics and overcomes the first model's restrictions on applicability.

Both approaches abstract away from real time progress and randomize choices in *logical* time. On the other hand, the relative speed - and thus, indirectly, the real-time behavior of the system's local processes - are crucial factors determining the outcome of probabilistic choices, even if non-determinism is absent from the system.

In another line of research [44] we have studied the likelihood of occurrence of non-sequential runs under random durations in a stochastic Petri net setting. It remains to better understand the properties of the probability measures thus obtained, to relate them with the models in logical time, and exploit them e.g. in *diagnosis*.

3.2.2.2. Distributed Markov Decision Processes

Participant: Serge Haddad.

Distributed systems featuring non-deterministic and probabilistic aspects are usually hard to analyze and, more specifically, to optimize. Furthermore, high complexity theoretical lower bounds have been established for models like partially observed Markovian decision processes and distributed partially observed Markovian decision processes. We believe that these negative results are consequences of the choice of the models rather than the intrinsic complexity of problems to be solved. Thus we plan to introduce new models in which the associated optimization problems can be solved in a more efficient way. More precisely, we start by studying connection protocols weighted by costs and we look for online and offline strategies for optimizing the mean cost to achieve the protocol. We have been cooperating on this subject with the SUMO team at Inria Rennes; in the joint work [36]; there, we strive to synthesize for a given MDP a control so as to guarantee a specific stationary behavior, rather than - as is usually done - so as to maximize some reward.

3.2.3. Large scale probabilistic systems

Addressing large-scale probabilistic systems requires to face state explosion, due to both the discrete part and the probabilistic part of the model. In order to deal with such systems, different approaches have been proposed:

- Restricting the synchronization between the components as in queuing networks allows to express the steady-state distribution of the model by an analytical formula called a product-form [41].

- Some methods that tackle with the combinatory explosion for discrete-event systems can be generalized to stochastic systems using an appropriate theory. For instance symmetry based methods have been generalized to stochastic systems with the help of aggregation theory [46].
- At last simulation, which works as soon as a stochastic operational semantic is defined, has been adapted to perform statistical model checking. Roughly speaking, it consists to produce a confidence interval for the probability that a random path fulfills a formula of some temporal logic [58] .

We want to contribute to these three axes: (1) we are looking for product-forms related to systems where synchronization are more involved (like in Petri nets [2]); (2) we want to adapt methods for discrete-event systems that require some theoretical developments in the stochastic framework and, (3) we plan to address some important limitations of statistical model checking like the expressiveness of the associated logic and the handling of rare events.

3.2.4. Real time distributed systems

Nowadays, software systems largely depend on complex timing constraints and usually consist of many interacting local components. Among them, railway crossings, traffic control units, mobile phones, computer servers, and many more safety-critical systems are subject to particular quality standards. It is therefore becoming increasingly important to look at networks of timed systems, which allow real-time systems to operate in a distributed manner.

Timed automata are a well-studied formalism to describe reactive systems that come with timing constraints. For modeling distributed real-time systems, networks of timed automata have been considered, where the local clocks of the processes usually evolve at the same rate [54] [45]. It is, however, not always adequate to assume that distributed components of a system obey a global time. Actually, there is generally no reason to assume that different timed systems in the networks refer to the same time or evolve at the same rate. Any component is rather determined by local influences such as temperature and workload.

3.2.4.1. Implementation of Real-Time Concurrent Systems

Participants: Thomas Chatain, Stefan Haar, Serge Haddad.

This was one of the tasks of the ANR ImpRo.

Formal models for real-time systems, like timed automata and time Petri nets, have been extensively studied and have proved their interest for the verification of real-time systems. On the other hand, the question of using these models as specifications for designing real-time systems raises some difficulties. One of those comes from the fact that the real-time constraints introduce some artifacts and because of them some syntactically correct models have a formal semantics that is clearly unrealistic. One famous situation is the case of Zeno executions, where the formal semantics allows the system to do infinitely many actions in finite time. But there are other problems, and some of them are related to the distributed nature of the system. These are the ones we address here.

One approach to implementability problems is to formalize either syntactical or behavioral requirements about what should be considered as a reasonable model, and reject other models. Another approach is to adapt the formal semantics such that only realistic behaviors are considered.

These techniques are preliminaries for dealing with the problem of implementability of models. Indeed implementing a model may be possible at the cost of some transformation, which make it suitable for the target device. By the way these transformations may be of interest for the designer who can now use high-level features in a model of a system or protocol, and rely on the transformation to make it implementable.

We aim at formalizing and automating translations that preserve both the timed semantics and the concurrent semantics. This effort is crucial for extending concurrency-oriented methods for logical time, in particular for exploiting partial order properties. In fact, validation and management - in a broad sense - of distributed systems is not realistic *in general* without understanding and control of their real-time dependent features; the link between real-time and logical-time behaviors is thus crucial for many aspects of *MExICo*'s work.

4. Application Domains

4.1. Telecommunications

Participants: Stefan Haar, Serge Haddad.

Stefan Haar, Serge Haddad.

MExICo's research is motivated by problems of system management in several domains, such as:

- In the domain of service oriented computing, it is often necessary to insert some Web service into an existing orchestrated business process, e.g. to replace another component after failures. This requires to ensure, often actively, conformance to the interaction protocol. One therefore needs to synthesize adaptators for every component in order to steer its interaction with the surrounding processes.
- Still in the domain of telecommunications, the supervision of a network tends to move from out-of-band technology, with a fixed dedicated supervision infrastructure, to in-band supervision where the supervision process uses the supervised network itself. This new setting requires to revisit the existing supervision techniques using control and diagnosis tools.

Currently, we have no active cooperation on these subjects.

4.2. Biological Systems

Participants: Thomas Chatain, Matthias Függer, Stefan Haar, Serge Haddad, Stefan Schwoon.

We have begun in 2014 to examine concurrency issues in systems biology, and are currently enlarging the scope of our research's applications in this direction. To see the context, note that in recent years, a considerable shift of biologists' interest can be observed, from the mapping of static genotypes to gene expression, i.e. the processes in which genetic information is used in producing functional products. These processes are far from being uniquely determined by the gene itself, or even jointly with static properties of the environment; rather, regulation occurs throughout the expression processes, with specific mechanisms increasing or decreasing the production of various products, and thus modulating the outcome. These regulations are central in understanding cell fate (how does the cell differenciate ? Do mutations occur ? etc), and progress there hinges on our capacity to analyse, predict, monitor and control complex and variegated processes. We have applied Petri net unfolding techniques for the efficient computation of attractors in a regulatory network; that is, to identify strongly connected reachability components that correspond to stable evolutions, e.g. of a cell that differentiates into a specific functionality (or mutation). This constitutes the starting point of a broader research with Petri net unfolding techniques in regulation. In fact, the use of ordinary Petri nets for capturing regulatory network (RN) dynamics overcomes the limitations of traditional RN models : those impose e.g. Monotonicity properties in the influence that one factor had upon another, i.e. always increasing or always decreasing, and were thus unable to cover all actual behaviours. Rather, we follow the more refined model of boolean networks of automata, where the local states of the different factors jointly detemine which state transitions are possible. For these connectors, ordinary PNs constitute a first approximation, improving greatly over the literature but leaving room for improvement in terms of introducing more refined logical connectors. Future work thus involves transcending this class of PN models. Via unfoldings, one has access – provided efficient techniques are available – to all behaviours of the model, rather than over-or under-approximations as previously. This opens the way to efficiently searching in particular for determinants of the cell fate : which attractors are reachable from a given stage, and what are the factors that decide in favor of one or the other attractor, etc. Our current research focusses cellular reprogramming on the one hand, and **distributed algorithms in wild or synthetic biological systems** on the other.

The latter is a distributed algorithms' view on microbiological systems, both with the goal to model and analyze existing microbiological systems as distributed systems, and to design and implement distributed algorithms in synthesized microbiological systems. Envisioned major long-term goals are drug production and medical treatment via synthesized bacterial colonies. We are approaching our goal of a distributed algorithm's view of microbiological systems from several directions: (i) Timing plays a crucial role in microbiological systems. Similar to modern VLSI circuits, dominating loading effects and noise render classical delay models unfeasible. In previous work we showed limitations of current delay models and presented a class of new delay models, so called involution channels. In [26] we showed that involution channels are still in accordance with Newtonian physics, even in presence of noise. (ii) In [7] we analyzed metastability in circuits by a three-valued Kleene logic, presented a general technique to build circuits that can tolerate a certain degree of metastability at its inputs, and showed the presence of a computational hierarchy. Again, we expect metastability to play a crucial role in microbiological systems, as similar to modern VLSI circuits, loading effects are pronounced. (iii) We studied agreement problems in highly dynamic networks without stability guarantees [28], [27]. We expect such networks to occur in bacterial cultures where bacteria communicate by producing and sensing small signal molecules like AHL. Both works also have theoretically relevant implications: The work in [27] presents the first approximate agreement protocol in a multidimensional space with time complexity independent of the dimension, working also in presence of Byzantine faults. In [28] we proved a tight lower bound on convergence rates and time complexity of asymptotic and approximate agreement in dynamic and classical static fault models. (iv) We are currently working with Da-Jung Cho, Manish Kushwaha (INRA), and Thomas Nowak (LRI) on biological infection models for *E. coli* colonies and M13 phages.

4.3. Autonomous Vehicles

Participant: Serge Haddad.

The validation of safety properties is a crucial concern for the design of computer guided systems, in particular for automated transport systems Our approach consists in analyzing the interactions of a randomized environment (roads, cross-sections, etc.) with a vehicle controller.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Reaching agreement in unstable times

Reaching approximate agreement in a distributed system among a set of local input values is a problem that often is repeatedly solved in artificial and natural distributed systems. Time efficient algorithms for this problem are thus of great theoretical and practical relevance. In [28] we studied the performance of such algorithms in dynamic networks. We showed lower time complexity bounds, demonstrating that already relatively simple broadcast and averaging algorithms achieve optimal time complexity. The results also imply new tight lower time complexity bounds for approximate agreement in classic distributed computing models with stable network architectures; solving a previously open problem.

5.1.2. New Semantics and State Spaces for Biological networks (and beyond)

We have gained major new insights into the dynamics of biological networks by

- obtaining [34], on the one hand, bi-directional translations between Contextual nets and BNs and correspondences between results on synchronism sensitivities. Taking advantage of CPN semantics enabling more behaviour than the generalized asynchronous updating mode, we propose an encoding of BNs that ensures correct abstraction of any multivalued refinement; and

- [20], [32] investigating update modes for discrete networks. It is commonly expected that Boolean networks produce an over-approximation of behaviours (reachable configurations), and that subsequent refinements would only prune some impossible transitions. However, we show that even generalized asynchronous updating of Boolean networks, which subsumes the usual updating modes including synchronous and fully asynchronous, does not capture all transitions doable in a multi-valued or timed refinement. We introduce a new semantics for interpreting BNs which meets with a correct abstraction of any multivalued refinements, with any update mode. This semantics subsumes all the usual updating modes, while enabling new behaviours achievable by more concrete models. Moreover, it appears that classical dynamical analyses of reachability and attractors have a simpler computational complexity: – reachability can be assessed in a polynomial number of iterations (instead of being PSPACE-complete with update modes); – attractors are hypercubes, and deciding the existence of attractors with a given upper-bounded dimension is in NP (instead of PSPACE-complete with update modes). The computation of iterations is in NP in the very general case, and is linear when local functions are monotonic, or with some usual representations of functions of BNs (binary decision diagrams, Petri nets, automata networks, etc.). In brief, the most permissive semantics of BNs enables a correct abstract reasoning on dynamics of BNs, with a greater tractability than previously introduced update modes. These works open new perspectives in concurrent semantics, and at the same time will allow to capture hitherto inaccessible phenotypes and pathways in biological networks.

5.1.3. Awards

-

6. New Software and Platforms

6.1. COSMOS

KEYWORD: Model Checker
FUNCTIONAL DESCRIPTION: COSMOS is a statistical model checker for the Hybrid Automata Stochastic Logic (HASL). HASL employs Linear Hybrid Automata (LHA), a generalization of Deterministic Timed Automata (DTA), to describe accepting execution paths of a Discrete Event Stochastic Process (DESP), a class of stochastic models which includes, but is not limited to, Markov chains. As a result HASL verification turns out to be a unifying framework where sophisticated temporal reasoning is naturally blended with elaborate reward-based analysis. COSMOS takes as input a DESP (described in terms of a Generalized Stochastic Petri Net), an LHA and an expression Z representing the quantity to be estimated. It returns a confidence interval estimation of Z, recently, it has been equipped with functionalities for rare event analysis.

It is easy to generate and use a C code for discrete Simulink models (using only discrete blocks, which are sampled at fixed intervals) using MathWorks tools. However, it limits the expressivity of the models. In order to use more diverse Simulink models and control the flow of a multi-model simulation (with Discrete Event Stochastic Processes) we developed a Simulink Simulation Engine embedded into Cosmos.

COSMOS is written in C++

- Participants: Benoît Barbot, Hilal Djafri, Marie Duflot-Kremer, Paolo Ballarini and Serge Haddad
- Contact: Benoît Barbot
- URL: http://www.lsv.ens-cachan.fr/~barbot/cosmos/

6.2. CosyVerif

FUNCTIONAL DESCRIPTION: CosyVerif is a platform dedicated to the formal specification and verification of dynamic systems. It allows to specify systems using several formalisms (such as automata and Petri nets), and to run verification tools on these models.

- Participants: Alban Linard, Fabrice Kordon, Laure Petrucci and Serge Haddad
- Partners: LIP6 - LSV - LIPN (Laboratoire d'Informatique de l'Université Paris Nord)
- Contact: Serge Haddad
- URL: http://www.cosyverif.org/

6.3. Mole

FUNCTIONAL DESCRIPTION: Mole computes, given a safe Petri net, a finite prefix of its unfolding. It is designed to be compatible with other tools, such as PEP and the Model-Checking Kit, which are using the resulting unfolding for reachability checking and other analyses. The tool Mole arose out of earlier work on Petri nets.

- Participant: Stefan Schwoon
- Contact: Stefan Schwoon
- URL: http://www.lsv.ens-cachan.fr/~schwoon/tools/mole/

7. New Results

7.1. Contract Based Design of Symbolic Controllers for Interconnected Multiperiodic Sampled-Data Systems

This paper deals with the synthesis of symbolic controllers for interconnected sampled-data systems where each component has its own sampling period. A compositional approach based on continuous-time assume-guarantee contracts is used. We provide sufficient conditions guaranteeing for a sampled-data system, satisfaction of an assume-guarantee contract and completeness of trajectories. Then, compositional results can be used to reason about interconnection of mul-tiperiodic sampled-data systems. We then show how discrete abstractions and symbolic control techniques can be applied to enforce the satisfaction of contracts and ensure completeness of trajectories. Finally, theoretical results are applied to a vehicle platooning problem on a circular road, which show the effectiveness of our approach.

7.2. Boolean Networks: Beyond Generalized Asynchronicity

Boolean networks are commonly used in systems biology to model dynamics of biochemical networks by abstracting away many (and often unknown) parameters related to speed and species activity thresholds. It is then expected that Boolean networks produce an over-approximation of behaviours (reachable configurations), and that subsequent refinements would only prune some impossible transitions. However, we show that even generalized asynchronous updating of Boolean networks, which subsumes the usual updating modes including synchronous and fully asynchronous, does not capture all transitions doable in a multi-valued or timed refinement. We define a structural model transformation which takes a Boolean network as input and outputs a new Boolean network whose asynchronous updating simulates both synchronous and asynchronous updating of the original network, and exhibits even more behaviours than the generalized asynchronous updating. We argue that these new behaviours should not be ignored when analyzing Boolean networks, unless some knowledge about the characteristics of the system explicitly allows one to restrict its behaviour.

7.3. Most Permissive Semantics of Boolean Networks

The usual update modes of Boolean networks (BNs), including synchronous and (generalized) asynchronous, fail to capture behaviours introduced by multivalued refinements. Thus, update modes do not allow a correct abstract reasoning on dynamics of biological systems, as they may lead to reject valid BN models. We introduce a new semantics for interpreting BNs which meets with a correct abstraction of any multivalued refinements, with any update mode. This semantics subsumes all the usual updating modes, while enabling new behaviours achievable by more concrete models. Moreover, it appears that classical dynamical analyses of reachability and attractors have a simpler computational complexity: – reachability can be assessed in a polynomial number of iterations (instead of being PSPACE-complete with update modes); – attractors are hypercubes, and deciding the existence of attractors with a given upper-bounded dimension is in NP (instead of PSPACE-complete with update modes). The computation of iterations is in NP in the very general case, and is linear when local functions are monotonic, or with some usual representations of functions of BNs (binary decision diagrams, Petri nets, automata networks, etc.). In brief, the most permissive semantics of BNs enables a correct abstract reasoning on dynamics of BNs, with a greater tractability than previously introduced update modes. This technical report lists the main definitions and properties of the most permissive semantics of BNs, and draw some remaining open questions.

7.4. Concurrency in Boolean networks

Boolean networks (BNs) are widely used to model the qualitative dynamics of biological systems. Besides the logical rules determining the evolution of each component with respect to the state of its regulators, the scheduling of components updates can have a dramatic impact on the predicted behaviours. In this paper, we explore the use of Contextual Petri Nets (CPNs) to study dynamics of BNs with a concurrency theory perspective. After showing bi-directional translations between CPNs and BNs and analogies between results on synchronism sensitivies, we illustrate that usual updating modes for BNs can miss plausible behaviours, i.e., incorrectly conclude on the absence/impossibility of reaching specific configurations. Taking advantage of CPN semantics enabling more behaviour than the generalized asynchronous updating mode, we propose an encoding of BNs ensuring a correct abstraction of any multivalued refinement, as one may expect to achieve when modelling biological systems with no assumption on its time features.

7.5. On the Composition of Discrete and Continuous-time Assume-Guarantee Contracts for Invariance

Many techniques for verifying invariance properties are limited to systems of moderate size. In this paper, we propose an approach based on assume-guarantee contracts and compositional reasoning for verifying invariance properties of a broad class of discrete-time and continuous-time systems consisting of interconnected components. The notion of assume-guarantee contracts makes it possible to divide responsibilities among the system components: a contract specifies an invariance property that a component must fulfill under some assumptions on the behavior of its environment (i.e. of the other components). We define weak and strong semantics of assume-guarantee contracts for both discrete-time and continuous-time systems. We then establish a certain number of results for compositional reasoning, which allow us to show that a global invariance property of the whole system is satisfied when all components satisfy their own contract. Interestingly, we show that the weak satisfaction of the contract is sufficient to deal with cascade compositions, while strong satisfaction is needed to reason about feedback composition. Specific results for systems described by differential inclusions are then developed. Throughout the paper, the main results are illustrated using simple examples.

7.6. Compositional synthesis of state-dependent switching control

We present a correct-by-design method of state-dependent control synthesis for sampled switching systems. Given a target region R of the state space, our method builds a capture set S and a control that steers any element of S into R. The method works by iterated backward reachability from R. The method is also used to synthesize a recurrence control that makes any state of R return to R infinitely often. We explain how the synthesis method can be performed in a compositional manner, and apply it to the synthesis of a compositional control of a concrete floor-heating system with 11 rooms and up to $2^11 = 2048$ toswitching modes.

7.7. An Improved Algorithm for the Co3ntrol Synthesis of Nonlinear Sampled Switched Systems

A novel algorithm for the control synthesis for nonlinear switched systems is presented in this paper. Based on an existing procedure of state-space bisection and made available for nonlinear systems with the help of guaranteed integration, the algorithm has been improved to be able to consider longer patterns of modes with a better pruning approach. Moreover, the use of guaranteed integration also permits to take bounded perturbations and varying parameters into account. It is particularly interesting for safety critical applications, such as in aeronautical, military or medical fields. The whole approach is entirely guaranteed and the induced controllers are correct-by-design. Some experimentations are performed to show the important gain of the new algorithm.

7.8. Control Synthesis for Stochastic Switched Systems using the Tamed Euler Method

In this paper, we explain how, under the one-sided Lipschitz (OSL) hypothesis, one can find an error bound for a variant of the Euler-Maruyama approximation method for stochastic switched systems. We then explain how this bound can be used to control stochastic switched switched system in order to stabilize them in a given region. The method is illustrated on several examples of the literature.

7.9. The Complexity of Diagnosability and Opacity Verification for Petri Nets

Diagnosability and opacity are two well-studied problems in discrete-event systems. We revisit these two problems with respect to expressiveness and complexity issues. We first relate different notions of diagnosability and opacity. We consider in particular fairness issues and extend the definition of Germanos et al. [ACM TECS, 2015] of weakly fair diagnosability for safe Petri nets to general Petri nets and to opacity questions. Second, we provide a global picture of complexity results for the verification of diagnosability and opacity. We show that diagnosability is NL-complete for finite state systems, PSPACE-complete for safe Petri nets (even with fairness), and EXPSPACE-complete for general Petri nets without fairness, while non diagnosability is inter-reducible with reachability when fault events are not weakly fair. Opacity is ESPACE-complete for safe Petri nets (even with fairness) and undecidable for general Petri nets already without fairness.

7.10. Integrating Simulink Models into the Model Checker Cosmos

We present an implementation for Simulink model executions in the statistical model-checker Cosmos. We take profit of this implementation for an hybrid modeling combining Petri nets and Simulink models.,Nous présentons une implémentation pour l'exécution de modèles Simulink dans le model-checker Cosmos. Cette implémentation est ensuite utilisée pour la simulation de modèles hybrides, combinant des réseaux de Petri et des modèles Simulink.

7.11. Bounds Computation for Symmetric Nets

Monotonicity in Markov chains is the starting point for quantitative abstraction of complex probabilistic systems leading to (upper or lower) bounds for probabilities and mean values relevant to their analysis. While numerous case studies exist in the literature, there is no generic model for which monotonicity is directly derived from its structure. Here we propose such a model and formalize it as a subclass of Stochastic Symmetric (Petri) Nets (SSNs) called Stochastic Monotonic SNs (SMSNs). On this subclass the monotonicity is proven by coupling arguments that can be applied on an abstract description of the state (symbolic marking). Our class includes both process synchronizations and resource sharings and can be extended to model open or cyclic closed systems. Automatic methods for transforming a non monotonic system into a monotonic one matching the MSN pattern, or for transforming a monotonic system with large state space into one with reduced state space are presented. We illustrate the interest of the proposed method by expressing standard monotonic models and modelling a flexible manufacturing system case study.

7.12. Distributed computation of vector clocks in Petri nets unfolding for test selection

Petri net unfoldings with time stamps allow to build distributed testers for distributed systems. However, the construction of the annotated unfolding of a distributed system currently remains a centralized task. In the aforemention paper, we extend a distributed unfolding technique in order to annotate the resulting unfolding with time stamps. This allows for distributed construction of distributed testers for distributed systems.

7.13. Hyper Partial Order Logic

We define HyPOL, a local hyper logic for partial order models, expressing properties of sets of runs. These properties depict shapes of causal dependencies in sets of partially ordered executions,with similarity relations defined as isomorphisms of past observations. Unsurprisingly, since comparison of projections are included, satisfiability of this logic is undecidable. We then addressmodel checking of HyPOL and show that, already for safe Petri nets, the problem is undecidable. Fortunately, sensible restrictions of observations and nets allow us to bring back model checking ofHyPOL to a decidable problem, namely model checking of MSO on graphs of bounded treewidth.

7.14. Integrating Simulink Models into the Model Checker Cosmos

We present an implementation for Simulink model executions in the statistical model-checker Cosmos. We take profit of this implementation for hybrid modeling and simulations combining Petri nets and Simulink models.

7.15. Site-Directed Deletion

We introduce a new bio-inspired operation called a site-directed deletion motivated from site-directed mutagenesis performed by enzymatic activity of DNA polymerase: Given two strings x and y, a site-directed deletion partially deletes a substring of x guided by the string y that specifies which part of a substring can be deleted. We study a few decision problems with respect to the new operation and examine the closure properties of the (iterated) site-directed deletion operations. We, then, define a site-directed deletion-closed (and-free) language L and investigate its decidability properties when L is regular or context-free.

7.16. Site-Directed Insertion: Decision Problems, Maximality and Minimality

Site-directed insertion is an overlapping insertion operation that can be viewed as analogous to the overlap assembly or chop operations that concatenate strings by overlapping a suffix and a prefix of the argument strings. We consider decision problems and language equations involving site-directed insertion. By relying on the tools provided by semantic shuffle on trajectories we show that one variable equations involving site-directed insertion and regular constants can be solved. We consider also maximal and minimal variants of the site-directed insertion operation.

7.17. A Faithful Binary Circuit Model with Adversarial Noise

Accurate delay models are important for static and dynamic timing analysis of digital circuits, and mandatory for formal verification. However, Függer et al. [IEEE TC 2016] proved that pure and inertial delays, which are employed for dynamic timing analysis in state-of-the-art tools like ModelSim, NC-Sim and VCS, do not yield faithful digital circuit models. Involution delays, which are based on delay functions that are mathematical involutions depending on the previous-output-to-input time offset, were introduced by Függer et al. [DATE'15] as a faithful alternative (that can easily be used with existing tools). Although involution delays were shown to predict real signal traces reasonably accurately, any model with a deterministic delay function is naturally limited in its modeling power. In this paper, we thus extend the involution model, by adding non-deterministic delay variations (random or even adversarial), and prove analytically that faithfulness is not impaired by this generalization. Albeit the amount of non-determinism must be considerably restricted to ensure this property, the result is surprising: the involution model differs from non-faithful models mainly in handling fast glitch trains, where small delay shifts have large effects. This originally suggested that adding even small variations should break the faithfulness of the model, which turned out not to be the case. Moreover, the results of our simulations also confirm that this generalized involution model has larger modeling power and, hence, applicability.

7.18. Tight Bounds for Asymptotic and Approximate Consensus

We study the performance of asymptotic and approximate consensus algorithms under harsh environmental conditions. The asymptotic consensus problem requires a set of agents to repeatedly set their outputs such that the outputs converge to a common value within the convex hull of initial values. This problem, and the related approximate consensus problem, are fundamental building blocks in distributed systems where exact consensus among agents is not required or possible, e.g., man-made distributed control systems , and have applications in the analysis of natural distributed systems, such as flocking and opinion dynamics. We prove tight lower bounds on the contraction rates of asymptotic consensus algorithms in dynamic networks, from which we deduce bounds on the time complexity of approximate consensus algorithms. In particular, the obtained bounds show optimality of asymptotic and approximate consensus algorithms presented in [Charron-Bost et al., ICALP'16] for certain dynamic networks, including the weakest dynamic network model in which asymptotic and approximate consensus are solvable. As a corollary we also obtain asymptotically tight bounds for asymptotic consensus in the classical asynchronous model with crashes. Central to our lower bound proofs is an extended notion of valency, the set of reachable limits of an asymptotic consensus algorithm starting from a given configuration. We further relate topological properties of valencies to the solvability of exact consensus , shedding some light on the relation of these three fundamental problems in dynamic networks.

7.19. Pomsets and Unfolding of Reset Petri Nets

Reset Petri nets are a particular class of Petri nets where transition firings can remove all tokens from a place without checking if this place actually holds tokens or not. In this paper we look at partial order semantics of such nets. In particular, we propose a pomset bisimulation for comparing their concurrent behaviours. Building on this pomset bisimulation we then propose a generalization of the standard finite complete prefixes of unfolding to the class of safe reset Petri nets.

7.20. Fast All-Digital Clock Frequency Adaptation Circuit for Voltage Droop Tolerance

Naive handling of supply voltage droops in synchronous circuits results in conservative bounds on clock speeds, resulting in poor performance even if droops are rare. Adaptive strategies detect such potentially hazardous events and either initiate a rollback to a previous state or proactively reduce clock speed in order to prevent timing violations. The performance of such solutions critically depends on a very fast response to droops. However, state-of-the-art solutions incur synchronization delay to avoid that the clock signal is affected by metastability. Addressing the challenges discussed by Keith Bowman in his ASYNC 2017 keynote talk, we present an all-digital circuit that can respond to droops within a fraction of a clock cycle. This is achieved by delaying clock signals based on measurement values while they undergo synchronization simultaneously. We verify our solution by formally proving correctness, complemented by VHDL and Spice simulations of a 65 nm ASIC design confirming the theoretically obtained results.

7.21. Fast Multidimensional Asymptotic and Approximate Consensus

We study the problems of asymptotic and approximate consensus in which agents have to get their values arbitrarily close to each others' inside the convex hull of initial values, either without or with an explicit decision by the agents. In particular, we are concerned with the case of multidimensional data, i.e., the agents' values are d-dimensional vectors. We introduce two new algorithms for dynamic networks, subsuming classical failure models like asynchronous message passing systems with Byzantine agents. The algorithms are the first to have a contraction rate and time complexity independent of the dimension d. In particular, we improve the time complexity from the previously fastest approximate consensus algorithm in asynchronous message passing systems with Byzantine faults by Mendes et al. [Distrib. Comput. 28].

7.22. Parameter Space Abstraction and Unfolding Semantics of Discrete Regulatory Networks

The modelling of discrete regulatory networks combines a graph specifying the pairwise influences between the variables of the system, and a parametrisation from which can be derived a discrete transition system. Given the influence graph only, the exploration of admissible parametrisations and the behaviours they enable is computationally demanding due to the combinatorial explosions of both parametrisation and reachable state space. This article introduces an abstraction of the parametrisation space and its refinement to account for the existence of given transitions, and for constraints on the sign and observability of influences. The abstraction uses a convex sub-lattice containing the concrete parametrisation space specified by its infimum and supremum parametrisations. It is shown that the computed abstractions are optimal, i.e., no smaller convex sublattice exists. Although the abstraction may introduce over-approximation, it has been proven to be conservative with respect to reachability of states. Then, an unfolding semantics for Parametric Regulatory Networks is defined, taking advantage of concurrency between transitions to provide a compact representation of reachable transitions. A prototype implementation is provided: it has been applied to several examples of Boolean and multi-valued networks, showing its tractability for networks with numerous components.

7.23. Interval Iteration Algorithm for MDPs and IMDPs

Markov Decision Processes (MDP) are a widely used model including both non-deterministic and probabilistic choices. Minimal and maximal probabilities to reach a target set of states, with respect to a policy resolving non-determinism, may be computed by several methods including value iteration. This algorithm, easy to implement and efficient in terms of space complexity, iteratively computes the probabilities of paths of increasing length. However, it raises three issues: (1) defining a stopping criterion ensuring a bound on the approximation, (2) analysing the rate of convergence, and (3) specifying an additional procedure to obtain the exact values once a sufficient number of iterations has been performed. The first two issues are still open and, for the third one, an upper bound on the number of iterations has been proposed. Based on a graph analysis and transformation of MDPs, we address these problems. First we introduce an interval iteration algorithm, for which the stopping criterion is straightforward. Then we exhibit its convergence rate. Finally we significantly improve the upper bound on the number of iterations required to get the exact values. We extend our approach to also deal with Interval Markov Decision Processes (IMDP) that can be seen as symbolic representations of MDPs.

7.24. Diagnosability of Repairable Faults

The diagnosis problem for discrete event systems consists in deciding whether some fault event occurred or not in the system, given partial observations on the run of that system. Diagnosability checks whether a correct diagnosis can be issued in bounded time after a fault, for all faulty runs of that system. This problem appeared two decades ago and numerous facets of it have been explored, mostly for permanent faults. It is known for example that diagnosability of a system can be checked in polynomial time, while the construction of a diagnoser is exponential. The present paper examines the case of transient faults, that can appear and be repaired. Diagnosability in this setting means that the occurrence of a fault should always be detected in bounded time, but also before the fault is repaired. Checking this notion of diagnosability is proved to be PSPACE-complete. It is also shown that faults can be reliably counted provided the system is diagnosable for faults and for repairs.

7.25. Metastability-Containing Circuits

In digital circuits, metastability can cause deteriorated signals that neither are logical 0 nor logical 1, breaking the abstraction of Boolean logic. Synchronizers, the only traditional countermeasure, exponentially decrease the odds of maintained metastability over time. We propose a fundamentally different approach: It is possible to deterministically contain metastability by fine-grained logical masking so that it cannot infect the entire circuit. At the heart of our approach lies a time-and value-discrete model for metastability in synchronous clocked digital circuits, in which metastability is propagated in a worst-case fashion. The proposed model permits positive results and passes the test of reproducing Marino's impossibility results. We fully classify which functions can be computed by circuits with standard registers. Regarding masking registers, we show that more functions become computable with each clock cycle, and that masking registers permit exponentially smaller circuits for some tasks. Demonstrating the applicability of our approach, we present the first fault-tolerant distributed clock synchronization algorithm that deterministically guarantees correct behavior in the presence of metastability. As a consequence, clock domains can be synchronized without using synchronizers, enabling metastability-free communication between them.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Our cooperation with industry took place in the context of a multi-lateral SystemX project, see below.

9. Partnerships and Cooperations

9.1. Regional Initiatives

- Serge Haddad and Yann Duplouy have been participating in the *Simulation pour la sécurité du véhicule autonome*(SVA) project at SystemX, in cooperation with Renault, on the application of formal methods to the development of embedded systems for autonomous vehicles.
- Matthias Függer co-leads the Digicosme working group HicDiesMeus on "Highly Constrained Discrete Agents for Modeling Natural Systems" (parsys.lri.fr/HicDiesMeus).
- Matthias Függer participates in the Farman project Dicimus in collaboration with Thomas Nowak (LRI). The project is on modeling of bacterial interactions using techniques from distributed computing theory and VLSI design.

9.2. National Initiatives

- Thomas Chatain, Stefan Haar, Serge Haddad and Stefan Schwoon are participating in the ANR Project ALGORECELL.
- Matthias Függer participates in the ANR project FREDDA on verification and synthesis of distributed algorithms.
- Laurent Fribourg participates in Digicosme Emergence Project "CODECSY" in collaboration with Antoine Girard (CentraleSupelec).

9.3. European Initiatives

Serge Haddad is a member of the European project ERC EQualIS "Enhancing the Quality of Interacting Systems" headed by Patricia Bouyer.

9.4. International Initiatives

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

9.4.1.1. LifeForm

Title: Life Sciences need formal Methods !

International Partner (Institution - Laboratory - Researcher):

Newcastle University (United Kingdom) - School of Computing Science - Victor Khomenko

Start year: 2016

See also: http://projects.lsv.ens-cachan.fr/LifeForm/

This project extends an existing cooperation between the MEXICO team and Newcastle University on partial-order based formal methods for concurrent systems. We enlarge the partnership to bioinformatics and synthetic biology. The proposal addresses addresses challenges concerning formal specification, verification, monitoring and control of synthetic biological systems, with use cases conducted in the Center for Synthetic Biology and the Bioeconomy (CSBB) in Newcastle. A main challenge is to create a solid modelling framework based on Petri-net type models that allow for causality analysis and rapid state space exploration for verification, monitoring and control purposes; a potential extension to be investigated concerns the study of attractors and cell reprogramming in Systems Biology.

9.4.2. Inria International Partners

9.4.2.1. Informal International Partners

Josep Carmona (UPC Barcelona) visited us in April and July 2018. He collaborated with Thomas Chatain on process mining.

9.5. International Research Visitors

9.5.1. Visits to International Teams

9.5.1.1. Research Stays Abroad

- Juraj Kolčák has started, in August 2018, a 6-month research visit in the MMM group / NII Tokyo (Japan), funded by the ERATO project, to work with the PI, Prof. Ichiro Hasuo. Stefan Haar has visited that group from Oct 29 to Friday Nov 2, preceded by a visit to Prof. Tatsuya Akutsu's group at Kyoto University (Uji campus) on Oct 26.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

Matthias Függer was

- general co-chair of the IEEE International Symposium on Asynchronous Circuits and Systems (ASYNC) 2018 (www.async2018.wien)
- general co-chair of the Workshop of Emergent Algorithms and Network Dynamics (WENDY) 2018 (wendy.paris)

Serge Haddad is a member of the steering commitee of the Petri Nets conference.

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

Thomas Chatain was co-chair of the program committe of ACSD 2018 interes.institute/acsd2018/.

10.1.2.2. Member of the Conference Program Committees

- Matthias Függer was a PC member of
  - the *21st IEEE International Symposium on Design and Diagnostics of Electronic Circuits and Systems(DDECS 2018)*,
- Stefan Haar was a PC member of
  - the *18th International Conference on Applications of Concurrency to Systems Design(ACSD 2018)*,
  - the workshop *Algorithms and Theories for the Analysis of Event Data 2018(ATAED 2018)*, and
  - the *International Workshop on Petri Nets and Modeling(PeMod '18)*.
- Laurent Fribourg was a PC member of
  - *Model-Based Design of Cyber Physical Systems* (CyPhy'18), October 4-5, 2018, Torino, Italy,
  - *12th International Conference on Reachability Problems* (RP'18), September 24-26 2018, Marseille, France,
  - *8th International Conference on New Computational Methods for Inverse Problems* (NCMIP'18), Ecole normale supérieure Paris-Saclay, France.
- Serge Haddad was a PC member of
  - *12th International Workshop on Verification and Evaluation of Computer and Communication Systems (VECOS 2018)*, Grenoble, France, September 2018
  - *5th International Symposium on Formal Approaches to Parallel and Distributed Systems (4PAD 2018)*, Orléans, France, July 2018

10.1.2.3. Reviewer

- Matthias Függer reviewed for Automatica, ASYNC'18, DDECS'18, Philisophical Transactions, DISC'18, PODC'18, SIROCCO'18, STACS'18.
- Stefan Schwoon reviewed for MFCS and FSTTCS.
- Stefan Haar reviewed for FOSSACS 2019.

10.1.3. Journals

10.1.3.1. Member of the Editorial Boards

- Matthias Függer is guest editor for the special issue *Selected Papers from the 24th IEEE International Symposium on Asynchronous Circuits and Systems - ASYNC 2018*
- Stefan Haar is an associate editor for *Journal of Discrete Event Dynamic Systems*

10.1.3.2. Reviewer - Reviewing Activities

- Thomas Chatain reviewed for *Journal of Discrete Event Dynamic Systems*, *Transactions of the Society for Modeling and Simulation International*.
- Stefan Schwoon reviewed for *Journal of Discrete Event Dynamic Systems*, *Acta Informatica*, *ACM Transactions on Programming Languages and Systems*.
- Stefan Haar reviewed for *Scientific Annals of Computer Science* and *IEEE Transactions on Automatic Control*.

10.1.4. Invited Talks

- Serge Haddad gave an invited talk on “Time and Stochastic Petri Nets” at the tutorials of Petri Nets 2018, the 25th June 2018, Bratislava, Slovakia
- Matthias Függer gave an invited talk at ACSD’18 on "Challenges of circuit design: Circuits as robust distributed algorithms"

10.1.5. Scientific Expertise

- Stefan Schwoon acted as reviewer for the ERC 2018 Starting Grant call.
- Serge Haddad was expert for the allocation of the grants ”Prime d’Investissement Recherche de l’Université” of Sorbonne Université

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Note: we only list the teaching activities of researchers here, not those of our assistant and full professors.

Licence: Stefan Haar taught one half of the L3 course on formal languages (18 h EQ TD) at ENS Paris-Saclay.

Master: Matthias Függer and Stefan Haar each taught a module of 10 h EQTD in the *Jaques Herbrand* master MI course *Introduction à la recherche*.

Laurent Fribourg taught one half of M2 course on “Hybrid Automata” at MPRI (Master Parisien de Recherche en Informatique).

10.2.2. Supervision

PhD:

- Thomas Chatain is the supervisor of the PhD thesis of Mathilde Boltenhagen.
- Stefan Haar is the supervisor of the PhD theses of
  - **Juraj Kolčák** Unfoldings and Abstract Interpretation for Parametric Biological Regulatory Networks, started in March 2017, and of
  - **Hugues Mandon** on *Computational models and algorithms for the prediction of cell reprogramming strategies*, started on Oct. 1st,

  both at ENS Paris-Saclay.
- Laurent Fribourg is the supervisor of the PhD theses of
  - **Adnane Saoud***Compositional controller synthesis for cyber-physical systems*, started in October 2016, co-supervised by Antoine Girard (CentraleSupelec), funded by Digicosme projet Emergence Codecsys
  - **Jawher Jerray***Formal analysis of real-time systems*, started in October 2018, co-supervised by Etienne André (Paris 13), funded by University Paris 13, ED Galilée.
- Serge Haddad is the supervisor of the PhD thesis of Igor Khmelnitsky on Machine Learning and Verification of Infinite-state Systems co-supervised with Alain Finkel.

10.2.3. Juries

- Thomas Chatain reviewed the PhD thesis of Thi Thanh Huyen Nguyen, directed by Laure Petrucci and defended at Université Paris 13 in December 2018.
- Stefan Schwoon reviewed the PhD thesis of Adrien Pommellet, directed by Tayssir Touili and defended at Université Paris 13 in July 2018. He also acted as examinator for the PhD thesis of Huu Vu Nguyen.

10.3. Popularization

Laurent Fribourg was interviewed by *L’Édition de l’ Université Paris-Saclay* in: “La cyberphysique prépare l’usine de demain”, May 2018.

10.3.1. Internal or external Inria responsibilities

- Laurent Fribourg is Head of Institut Farman (FR 3311 CNRS & ENS Paris-Saclay).
- Serge Haddad is Head of the Computer Science Department of ENS Paris-Saclay.
- Stefan Haar is the president of Inria's COST-GTRI.

11. Bibliography

Major publications by the team in recent years

[1] S. HAAR, S. HADDAD, T. MELLITI, S. SCHWOON.*Optimal Constructions for Active Diagnosis*, in "Proceedings of the 33rd Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS'13)", Guwahati, India, A. SETH, N. VISHNOI (editors), Leibniz International Proceedings in Informatics, Leibniz-Zentrum für Informatik, December 2013, vol. 24, p. 527-539, http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/HHMS13-fsttcs.pdf

[2] S. HADDAD, J. MAIRESSE, H.-T. NGUYEN.*Synthesis and Analysis of Product-form Petri Nets*, in "Fundamenta Informaticae", 2013, vol. 122, n^{o} 1-2, p. 147-172, https://hal.archives-ouvertes.fr/hal-00925774

[3] A. LE COËNT, J. ALEXANDRE DIT SANDRETTO, A. CHAPOUTOT, L. FRIBOURG.*An Improved Algorithm for the Control Synthesis of Nonlinear Sampled Switched Systems*, in "Formal Methods in System Design", November 2017, vol. 53, n^{o} 3, p. 363-383 [*DOI :* 10.1007/s10703-017-0305-8], https://hal.archives-ouvertes.fr/hal-01399337

Publications of the year

Doctoral Dissertations and Habilitation Theses

[4] E. LEFAUCHEUX.*Controlling information in Probabilistic Systems*, Université Rennes 1, September 2018, https://hal.inria.fr/tel-01946840

Articles in International Peer-Reviewed Journal

[5] B. BÉRARD, S. HAAR, S. SCHMITZ, S. SCHWOON.*The Complexity of Diagnosability and Opacity Verification for Petri Nets*, in "Fundamenta Informaticae", July 2018, vol. 161, n^{o} 4, p. 317–349 [*DOI :* 10.3233/FI-2018-1706], https://hal.inria.fr/hal-01852119

[6] É. FABRE, L. HÉLOUËT, E. LEFAUCHEUX, H. MARCHAND.*Diagnosability of Repairable Faults*, in "Discrete Event Dynamic Systems", 2018, vol. 28, n^{o} 2, p. 183-213 [*DOI :* 10.1007/s10626-017-0255-8], https://hal.inria.fr/hal-01646911

[7] S. FRIEDRICHS, M. FÜGGER, C. LENZEN.*Metastability-Containing Circuits*, in "IEEE Transactions on Computers", March 2018, p. 1167 - 1183 [*DOI :* 10.1109/TC.2018.2808185], https://hal.inria.fr/hal-01936292

[8] P. GASTIN, B. MONMEGE.*A Unifying Survey on Weighted Logics and Weighted Automata: Core Weighted Logic: Minimal and Versatile Specification of Quantitative Properties*, in "Soft Computing", February 2018, vol. 22, n^{o} 4, p. 1047-1065 [*DOI :* 10.1007/s00500-015-1952-6], https://hal.archives-ouvertes.fr/hal-01130216

[9] S. HADDAD.*Memoryless Determinacy of Finite Parity Games: Another Simple Proof*, in "Information Processing Letters", 2018, n^{o} 132, 3, https://hal.inria.fr/hal-01541508

[10] S. HADDAD, B. MONMEGE.*Interval Iteration Algorithm for MDPs and IMDPs*, in "Theoretical Computer Science", July 2018, vol. 735, p. 111 - 131 [*DOI :* 10.1016/J.TCS.2016.12.003], https://hal.archives-ouvertes.fr/hal-01809094

[11] J. KOLČÁK, D. ŠAFRÁNEK, S. HAAR, L. PAULEVÉ.*Parameter Space Abstraction and Unfolding Semantics of Discrete Regulatory Networks*, in "Theoretical Computer Science", 2018 [*DOI :* 10.1016/J.TCS.2018.03.009], https://hal.archives-ouvertes.fr/hal-01734805

[12] F. KORDON, H. GARAVEL, L. HILLAH, E. PAVIOT-ADET, L. JEZEQUEL, F. HULIN-HUBARD, E. AMPARORE, M. BECCUTI, B. BERTHOMIEU, H. EVRARD, P. G. JENSEN, D. LE BOTLAN, T. LIEBKE, J. MEIJER, J. SRBA, Y. THIERRY-MIEG, J. VAN DE POL, K. WOLF.*MCC'2017-The Seventh Model Checking Contest*, in "LNCS Transactions on Petri Nets and Other Models of Concurrency ", 2018, vol. 11090, p. 181-209 [*DOI :* 10.1007/978-3-662-58381-4_9], https://hal.inria.fr/hal-01917492

[13] A. LE COËNT, L. FRIBOURG, N. MARKEY, F. DE VUYST, L. CHAMOIN.*Compositional synthesis of state-dependent switching control*, in "Theoretical Computer Science", November 2018, vol. 750, p. 53-68 [*DOI :* 10.1016/J.TCS.2018.01.021], https://hal-utc.archives-ouvertes.fr/hal-01958874

[14] A. LE COËNT, L. FRIBOURG, N. MARKEY, F. DE VUYST, L. CHAMOIN.*Compositional synthesis of state-dependent switching control*, in "Theoretical Computer Science", November 2018, vol. 750, p. 53-68 [*DOI :* 10.1016/J.TCS.2018.01.021], https://hal.archives-ouvertes.fr/hal-01860379

International Conferences with Proceedings

[15] B. BARBOT, B. BÉRARD, Y. DUPLOUY, S. HADDAD.*Integrating Simulink Models into the Model Checker Cosmos*, in "39th International Conference on Applications and Theory of Petri Nets and Concurrency", Bratislava, Slovakia, Lecture Notes in Computer Sciences, Springer, June 2018, vol. 10877, p. 363-373 [*DOI :* 10.1007/978-3-319-91268-4_19], https://hal.archives-ouvertes.fr/hal-01916467

[16] B. BÉRARD, S. HAAR, L. HELOUET.*Hyper Partial Order Logic*, in "FSTTCS 2018 - Foundations of Software Technology and Theoretical Computer Science", Ahmedabad, India, December 2018, https://hal.inria.fr/hal-01884390

[17] D.-J. CHO, Y.-S. HAN, H. KIM, K. SALOMAA.*Site-Directed Deletion*, in "22nd International Conference on Developments in Language Theory (DLT 2018)", Tokyo, Japan, September 2018, https://hal.inria.fr/hal-01937635

[18] D.-J. CHO, Y.-S. HAN, K. SALOMAA, T. J. SMITH.*Site-Directed Insertion: Decision Problems, Maximality and Minimality*, in "20th International Conference on Descriptional Complexity of Formal Systems", Halifax, Canada, July 2018, https://hal.inria.fr/hal-01937654

[19] T. CHATAIN, R. GROSU.*Message from the ACSD 2018 Program Chairs*, in "2018 18th International Conference on Application of Concurrency to System Design (ACSD)", Bratislava, France, IEEE, June 2018 [*DOI :* 10.1109/ACSD.2018.00005], https://hal.inria.fr/hal-01936815

[20] T. CHATAIN, S. HAAR, L. PAULEVÉ.*Boolean Networks: Beyond Generalized Asynchronicity*, in "AUTOMATA 2018 - 24th IFIP WG 1.5 International Workshop on Cellular Automata and Discrete Complex Systems", Ghent, Belgium, J. M. BAETENS, M. KUTRIB (editors), Lecture Notes in Computer Science, Springer, June 2018, vol. 10875, p. 29-42 [*DOI :* 10.1007/978-3-319-92675-9_3], https://hal.archives-ouvertes.fr/hal-01768359

[21] M. FÜGGER, A. KINALI, C. LENZEN, B. WIEDERHAKE.*Fast All-Digital Clock Frequency Adaptation Circuit for Voltage Droop Tolerance*, in "24th IEEE International Symposium on Asynchronous Circuits and Systems (ASYNC)", Wien, Austria, May 2018, https://hal.inria.fr/hal-01936403

[22] S. HAAR, J. KOLČÁK, L. PAULEVÉ.*Combining Refinement of Parametric Models with Goal-Oriented Reduction of Dynamics*, in "VMCAI 2019 - 20th International Conference on Verification, Model Checking, and Abstract Interpretation", Lisbon, Portugal, January 2019, https://arxiv.org/abs/1811.12377 , https://hal.archives-ouvertes.fr/hal-01940174

[23] L. JEZEQUEL, T. CHATAIN, M. COMLAN, D. DELFIEU, O. H. ROUX.*Pomsets and Unfolding of Reset Petri Nets*, in "12th International Conference on Language and Automata Theory and Applications (LATA 2018)", Ramat Gan, Israel, S. T. KLEIN, C. MARTÍN-VIDE, D. SHAPIRA (editors), Language and Automata Theory and Applications, Springer, April 2018, https://hal.archives-ouvertes.fr/hal-01766530

[24] L. JEZEQUEL, A. MADALINSKI, S. SCHWOON.*Distributed computation of vector clocks in Petri nets unfolding for test selection*, in "Workshop on Discrete Event Systems (WODES)", Sorrento, Italy, May 2018, https://hal.inria.fr/hal-01735406

[25] A. SAOUD, A. GIRARD, L. FRIBOURG.*Contract Based Design of Symbolic Controllers for Interconnected Multiperiodic Sampled-Data Systems*, in "57th IEEE Conference on Decision and Control (CDC 2018)", Miami Beach, FL, United States, December 2018, https://hal.archives-ouvertes.fr/hal-01857389

Conferences without Proceedings

[26] M. FÜGGER, J. MAIER, R. NAJVIRT, T. NOWAK, U. SCHMID.*A Faithful Binary Circuit Model with Adversarial Noise*, in "DATE 2018 - Design, Automation and Test in Europe Conference and Exhibition", Dresden, Germany, March 2018 [*DOI :* 10.23919/DATE.2018.8342219], https://hal.archives-ouvertes.fr/hal-01738254

[27] M. FÜGGER, T. NOWAK.*Fast Multidimensional Asymptotic and Approximate Consensus*, in "International Symposium on DIStributed Computing (DISC) 2018", New Orleans, United States, October 2018 [*DOI :* 10.4230/LIPIcs.DISC.2018.27], https://hal.archives-ouvertes.fr/hal-01936316

[28] M. FÜGGER, T. NOWAK, M. SCHWARZ.*Tight Bounds for Asymptotic and Approximate Consensus*, in "ACM Symposium on Principles of Distributed Computing (PODC'18)", Egham, United Kingdom, July 2018 [*DOI :* 10.1145/3212734.3212762], https://hal.archives-ouvertes.fr/hal-01799039

[29] A. LE COËNT, L. FRIBOURG, J. VACHER.*Control Synthesis for Stochastic Switched Systems using the Tamed Euler Method*, in "6th IFAC Conference on Analysis and Design of Hybrid Systems ADHS 2018 ADHS'18,", Oxford, United Kingdom, July 2018, vol. 51, n^{o} 16, https://hal.archives-ouvertes.fr/hal-01959845

[30] A. SAOUD, A. GIRARD, L. FRIBOURG.*On the Composition of Discrete and Continuous-time Assume-Guarantee Contracts for Invariance*, in "European Control Conference (ECC 2018)", Limassol, Cyprus, 2018, https://hal.archives-ouvertes.fr/hal-01712710

Research Reports

[31] B. BARBOT, B. BÉRARD, Y. DUPLOUY, S. HADDAD.*Integrating Simulink Models into the Model Checker Cosmos*, LSV, ENS Cachan, CNRS, Inria, Université Paris-Saclay, Cachan (France) ; LIP6, Sorbonne Université, CNRS, UMR 7606 ; LACL, Université Paris-Est, March 2018, https://hal.archives-ouvertes.fr/hal-01725835

[32] T. CHATAIN, S. HAAR, L. PAULEVÉ.*Most Permissive Semantics of Boolean Networks*, LRI, Univ. Paris-Sud, CNRS, Inria, Université Paris-Saclay ; LSV, ENS Cachan, CNRS, Inria, Université Paris-Saclay, Cachan (France), 2018, https://arxiv.org/abs/1808.10240 , https://hal.archives-ouvertes.fr/hal-01864693

Other Publications

[33] B. BARBOT, M. BECCUTI, G. FRANCESCHINIS, S. HADDAD, C. PICARONNY.*Bounds Computation for Symmetric Nets*, March 2018, working paper or preprint, https://hal.inria.fr/hal-01726011

[34] T. CHATAIN, S. HAAR, J. KOLČÁK, A. THAKKAR, L. PAULEVÉ.*Concurrency in Boolean networks*, October 2018, working paper or preprint, https://hal.inria.fr/hal-01893106

References in notes

[35] S. ABBES, A. BENVENISTE, S. HAAR.*A Petri net model for distributed estimation*, in "Proc. MTNS 2004, Sixteenth International Symposium on Mathematical Theory of Networks and Systems, Louvain (Belgium), ISBN 90-5682-517-8", 2004

[36] S. AKSHAY, N. BERTRAND, S. HADDAD, L. HELOUET.*The steady-state control problem for Markov decision processes*, in "Qest 2013", Buenos Aires, Argentina, K. R. JOSHI, M. SIEGLE, M. STOELINGA, P. R. D'ARGENIO (editors), Springer, September 2013, vol. 8054, p. 290-304, https://hal.inria.fr/hal-00879355

[37] R. ALUR, K. ETESSAMI, M. YANNAKAKIS.*Realizability and Verification of MSC Graphs*, in "Theor. Comput. Sci.", 2005, vol. 331, n^{o} 1, p. 97–114

[38] P. BALDAN, TH. CHATAIN, S. HAAR, B. KÖNIG.*Unfolding-based Diagnosis of Systems with an Evolving Topology*, in "Information and Computation", October 2010, vol. 208, n^{o} 10, p. 1169-1192, http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/BCHK-icomp10.pdf

[39] P. BALDAN, A. CORRADINI, B. KÖNIG, S. SCHWOON.*McMillan's complete prefix for contextual nets*, in "Transactions on Petri Nets and Other Models of Concurrency", November 2008, vol. 1, p. 199–220, Volume 5100 of Lecture Notes in Computer Science

[40] P. BALDAN, S. HAAR, B. KOENIG.*Distributed Unfolding of Petri Nets*, in "Proc.FOSSACS 2006", LNCS, Springer, 2006, vol. 3921, p. 126-141, Extended version: Technical Report CS-2006-1. Department of Computer Science, University Ca' Foscari of Venice

[41] F. BASKETT, K. M. CHANDY, R. R. MUNTZ, F. G. PALACIOS.*Open, Closed, and Mixed Networks of Queues with Different Classes of Customers*, in "J. ACM", April 1975, vol. 22, p. 248–260, http://doi.acm.org/10.1145/321879.321887

[42] A. BENVENISTE, É. FABRE, S. HAAR.*Markov Nets: Probabilistic Models for distributed and concurrent Systems*, in "IEEE Transactions on Automatic Control", 2003, vol. 48 (11), p. 1936-1950, Extended version: IRISA Research Report 1538

[43] P. BHATEJA, P. GASTIN, M. MUKUND, K. NARAYAN KUMAR.*Local testing of message sequence charts is difficult*, in "Proceedings of the 16th International Symposium on Fundamentals of Computation Theory (FCT'07)", Budapest, Hungary, E. CSUHAJ-VARJÚ, Z. ÉSIK (editors), Lecture Notes in Computer Science, Springer, August 2007, vol. 4639, p. 76-87 [*DOI :* 10.1007/978-3-540-74240-1_8], http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/BGMN-fct07.pdf

[44] A. BOUILLARD, S. HAAR, S. ROSARIO.*Critical paths in the Partial Order Unfolding of a Stochastic Petri Net*, in "Proceedings of the 7th International Conference on Formal Modelling and Analysis of Timed Systems (FORMATS'09)", Budapest, Hungary, J. OUAKNINE, F. VAANDRAGER (editors), Lecture Notes in Computer Science, Springer, September 2009, vol. 5813, p. 43-57 [*DOI :* 10.1007/978-3-642-04368-0_6], http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/BHR-formats09.pdf

[45] P. BOUYER, S. HADDAD, P.-A. REYNIER.*Timed Unfoldings for Networks of Timed Automata*, in "Proceedings of the 4th International Symposium on Automated Technology for Verification and Analysis (ATVA'06)", Beijing, ROC, S. GRAF, W. ZHANG (editors), Lecture Notes in Computer Science, Springer, October 2006, vol. 4218, p. 292-306, http://www.lsv.ens-cachan.fr/Publis/PAPERS/PDF/BHR-atva06.pdf

[46] G. CHIOLA, C. DUTHEILLET, G. FRANCESCHINIS, S. HADDAD.*Stochastic Well-Formed Colored Nets and Symmetric Modeling Applications*, in "IEEE Transactions on Computers", November 1993, vol. 42, n^{o} 11, p. 1343-1360, http://www.lsv.ens-cachan.fr/Publis/PAPERS/PS/CDFH-toc93.ps

[47] R. DEBOUK, D. TENEKETZIS.*Coordinated decentralized protocols for failure diagnosis of discrete-event systems*, in "Journal of Discrete Event Dynamical Systems: Theory and Application", 2000, vol. 10, p. 33–86

[48] J. ESPARZA, K. HELJANKO.*Unfoldings - A Partial-Order Approach to Model Checking*, EATCS Monographs in Theoretical Computer Science, Springer, 2008

[49] É. FABRE, A. BENVENISTE, C. JARD, S. HAAR.*Diagnosis of Asynchronous Discrete Event Systems, a Net Unfolding Approach*, in "IEEE Trans. Aut. Control", 2003, vol. 48 (5), p. 714-727

[50] É. FABRE, A. BENVENISTE, C. JARD, S. HAAR.*Distributed monitoring of concurrent and asynchronous systems*, in "Discrete Event Dynamic Systems: theory and application", 2005, vol. 15 (1), p. 33-84, Preliminary version: Proc. CONCUR 2003, LNCS 2761, pp.1–28, Springer

[51] S. HAAR, A. BENVENISTE, É. FABRE, C. JARD.*Partial Order Diagnosability Of Discrete Event Systems Using Petri Net Unfoldings*, in "42nd IEEE Conference on Decision and Control (CDC)", 2003

[52] S. HAAR.*Probabilistic Cluster Unfoldings*, in "Fundamenta Informaticae", 2003, vol. 53 (3-4), p. 281-314

[53] S. LAFORTUNE, Y. WANG, T.-S. YOO.*Diagnostic Decentralisé Des Systèmes A Evénements Discrets*, in "Journal Europeen des Systèmes Automatisés (RS-JESA)", August 2005, vol. 99, n^{o} 99, p. 95–110

[54] K. G. LARSEN, P. PETTERSSON, W. YI.*Compositional and symbolic model-checking of real-time systems*, in "Proc. of RTSS 1995", IEEE Computer Society, 1995, p. 76-89

[55] L. RICKER, K. RUDIE.*Know Means No: Incorporating Knowledge into Discrete-Event Control Systems*, in "IEEE Transactions on Automatic Control", September 2000, vol. 45, n^{o} 9, p. 1656–1668

[56] L. RICKER, K. RUDIE.*Knowledge Is a Terrible Thing to Waste: Using Inference in Discrete-Event Control Problems*, in "IEEE Transactions on Automatic Control", MarchSeptember 2007, vol. 52, n^{o} 3, p. 428–441

[57] C. RODRÍGUEZ, S. SCHWOON, V. KHOMENKO.*Contextual Merged Processes*, in "34th International Conference on Applications and Theory of Petri Nets (ICATPN'13)", Italy, Lecture Notes in Computer Science, Springer, 2013, vol. 7927, p. 29-48 [*DOI :* 10.1007/978-3-642-38697-8_3], https://hal.archives-ouvertes.fr/hal-00926202

[58] H. L. S. YOUNES, R. G. SIMMONS.*Statistical probabilistic model checking with a focus on time-bounded properties*, in "Inf. Comput.", September 2006, vol. 204, p. 1368–1409 [*DOI :* 10.1016/J.IC.2006.05.002], http://dl.acm.org/citation.cfm?id=1182767.1182770

Project-Team PARIETAL

Modelling brain structure, function and variability based on high-field MRI data.

IN COLLABORATION WITH: CEA Neurospin

IN PARTNERSHIP WITH:
Centre CEA-Saclay

RESEARCH CENTER
Saclay - Île-de-France

THEME
Computational Neuroscience and Medicine

Table of contents

Project-Team PARIETAL

Creation of the Project-Team: 2009 July 01

Keywords:

Computer Science and Digital Science:

A3.3. - Data and knowledge analysis
A3.3.2. - Data mining
A3.3.3. - Big data analysis
A3.4. - Machine learning and statistics
A3.4.1. - Supervised learning
A3.4.2. - Unsupervised learning
A3.4.4. - Optimization and learning
A3.4.5. - Bayesian methods
A3.4.6. - Neural networks
A3.4.7. - Kernel methods
A3.4.8. - Deep learning
A5.3.2. - Sparse modeling and image representation
A5.3.3. - Pattern recognition
A5.9.1. - Sampling, acquisition
A5.9.2. - Estimation, modeling
A5.9.6. - Optimization tools
A6.2.4. - Statistical methods
A6.2.6. - Optimization
A9.2. - Machine learning
A9.3. - Signal analysis

Other Research Topics and Application Domains:

B1.2. - Neuroscience and cognitive science
B1.2.1. - Understanding and simulation of the brain and the nervous system
B1.2.2. - Cognitive science
B2.2.6. - Neurodegenerative diseases
B2.6.1. - Brain imaging

1. Team, Visitors, External Collaborators

Research Scientists

Bertrand Thirion [Team leader, Inria, Senior Researcher, HDR]
Philippe Ciuciu [CEA, Researcher, HDR]
Denis Alexander Engemann [Inria, Starting Research Position]
Alexandre Gramfort [Inria, Senior Researcher, HDR]
Gaël Varoquaux [Inria, Researcher, HDR]
Demian Wassermann [Inria, Researcher]

External Collaborators

Sergul Aydore [Stevens Institute of Technology, until Apr 2018]

Alexandre Boucaud [CNRS, until Sep 2018]
Danilo Bzdok [Univ. Aachen, from Jul 2018]
Michael Eickenberg [Univ. Berkeley]
Mehdi Rahim [Inria, until Mar 2018]
Venkat Rajagoplan [Institut Telecom ex GET Groupe des Ecoles des Télécommunications , until Jun 2018]
Joseph Salmon [Institut Telecom ex GET Groupe des Ecoles des Télécommunications, HDR]
Moritz Boos [Univ. Oldenburg]

Technical Staff

Alexandre Abadie [Inria]
Jérémie Du Boisberranger [Inria, from Jun 2018]
Guillaume Favelier [Inria, from Dec 2018]
Ana Luisa Grilo Pinho [Inria]
Olivier Grisel [Inria]
Guillaume Lemaitre [Inria]
Joan Massich Vall [Inria]
Romuald Menuet [Inria, from Dec 2018]
Antonio Andre Monteiro Manoel [Inria, until Aug 2018]
Juan Jesus Torre Tresols [Inria, from Apr 2018]
Joris Van Den Bossche [Inria]
Kshitij Chawla [Inria, from Jul 2018]
Pierre Glaser [Inria, from Sep 2018]

PhD Students

Pierre Ablin [Inria]
Thomas Bazeille [Inria, from Oct 2018]
Quentin Bertrand [Inria, from Oct 2018]
Patricio Cerda Reyes [Inria]
Hamza Cherkaoui [CEA]
Jerome-Alexis Chevalier [Inria]
Kamalaker Reddy Dadi [Inria]
Jérome Dockès [Inria]
Loubna El Gueddari [Ministère de l'Enseignement Supérieur et de la Recherche]
Valentin Iovene [Inria, from Jul 2018]
Hubert Jacob Banville [Interaxon, from Sep 2018]
Maeliss Jallais [Inria, from Nov 2018]
Hicham Janati [Inria]
Sylvain Lanuzel [CEA, from Feb 2018]
Carole Lazarus [CEA, until Sep 2018]
Antonia Machlouzarides Shalit [Inria]
Mathurin Massias [Inria]
Arthur Mensch [Univ Paris-Sud, until Oct 2018]
Tuan Binh Nguyen [Inria, from Oct 2018]
Nicolas Prost [Ecole polytechnique, from Oct 2018]
Hugo Richard [Univ Paris-Saclay, from Sep 2018]
David Sabbagh [INSERM, from Oct 2018]

Post-Doctoral Fellows

Darya Chyzhyk [Inria, until Jul 2018]
Elvis Dohmatob [Inria, until Feb 2018]
Thomas Moreau [Inria, from Mar 2018]

Visiting Scientists

Lang Chen [Univ. Stanford, from Dec 2018]

Guillermo Gallardo [Univ Côte d'Azur, from Dec 2018]
Matt Higger [Harvard Medical School, from Dec 2018]

Administrative Assistant

Corinne Petitot [Inria]

2. Overall Objectives

2.1. Overall Objectives

The Parietal team focuses on mathematical methods for modeling and statistical inference based on neuroimaging data, with a particular interest in machine learning techniques and applications of human functional imaging. This general theme splits into four research axes:

- Modeling for neuroimaging population studies,
- Encoding and decoding models for cognitive imaging,
- Statistical and machine learning methods for large-scale data,
- Compressed-sensing for MRI.

Parietal is also strongly involved in open-source software development in scientific Python (machine learning) and for neuroimaging applications.

3. Research Program

3.1. Inverse problems in Neuroimaging

Many problems in neuroimaging can be framed as forward and inverse problems. For instance, brain population imaging is concerned with the *inverse problem* that consists in predicting individual information (behavior, phenotype) from neuroimaging data, while the corresponding *forward problem* boils down to explaining neuroimaging data with the behavioral variables. Solving these problems entails the definition of two terms: a loss that quantifies the goodness of fit of the solution (does the model explain the data well enough?), and a regularization scheme that represents a prior on the expected solution of the problem. These priors can be used to enforce some properties on the solutions, such as sparsity, smoothness or being piece-wise constant.

Let us detail the model used in typical inverse problem: Let $\mathbf{X}$ be a neuroimaging dataset as an $(n_{subjects}, n_{voxels})$ matrix, where $n_{subjects}$ and n_{voxels} are the number of subjects under study, and the image size respectively, $\mathbf{Y}$ a set of values that represent characteristics of interest in the observed population, written as $(n_{subjects}, n_{features})$ matrix, where $n_{features}$ is the number of characteristics that are tested, and $\mathbf{w}$ an array of shape $(n_{voxels}, n_{features})$ that represents a set of pattern-specific maps. In the first place, we may consider the columns $\mathbf{Y}_1, .., \mathbf{Y}_{n_{features}}$ of Y independently, yielding $n_{features}$ problems to be solved in parallel:

$$\mathbf{Y}_i = \mathbf{X}\mathbf{w}_i + \epsilon_i, \forall i \in \{1, .., n_{features}\},$$

where the vector contains $\mathbf{w}_i$ is the i^{th} row of $\mathbf{w}$. As the problem is clearly ill-posed, it is naturally handled in a regularized regression framework:

$$\widehat{w}_i = \text{argmin}_{w_i} \|\mathbf{Y}_i - \mathbf{X}\mathbf{w}_i\|^2 + \Psi(\mathbf{w}_i), \tag{10}$$

where Ψ is an adequate penalization used to regularize the solution:

$$\Psi(\mathbf{w}; \lambda_1, \lambda_2, \eta_1, \eta_2) = \lambda_1 \|\mathbf{w}\|_1 + \lambda_2 \|\mathbf{w}\|_2 + \eta_1 \|\nabla \mathbf{w}\|_{2,1} + \eta_2 \|\nabla \mathbf{w}\|_{2,2} \tag{11}$$

with λ_1, λ_2, η_1, $\eta_2 \geq 0$ (this formulation particularly highlights the fact that convex regularizers are norms or quasi-norms). In general, only one or two of these constraints is considered (hence is enforced with a non-zero coefficient):

- When $\lambda_1 > 0$ only (LASSO), and to some extent, when $\lambda_1, \lambda_2 > 0$ only (elastic net), the optimal solution $\mathbf{w}$ is (possibly very) sparse, but may not exhibit a proper image structure; it does not fit well with the intuitive concept of a brain map.
- Total Variation regularization (see Fig. 1) is obtained for ($\eta_1 > 0$ only), and typically yields a piece-wise constant solution. It can be associated with Lasso to enforce both sparsity and sparse variations.
- Smooth lasso is obtained with ($\eta_2 > 0$ and $\lambda_1 > 0$ only), and yields smooth, compactly supported spatial basis functions.

Note that, while the qualitative aspect of the solutions are very different, the predictive power of these models is often very close.

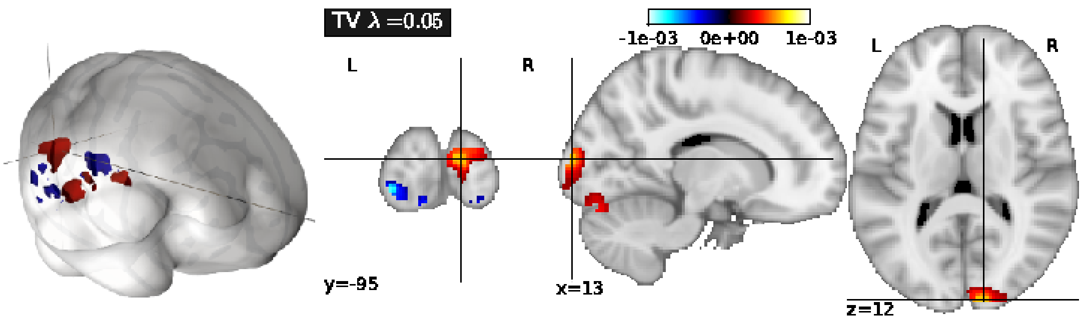


Figure 1. Example of the regularization of a brain map with total variation in an inverse problem. The problem here is to predict the spatial scale of an object presented as a stimulus, given functional neuroimaging data acquired during the presentation of an image. Learning and test are performed across individuals. Unlike other approaches, Total Variation regularization yields a sparse and well-localized solution that also enjoys high predictive accuracy.

The performance of the predictive model can simply be evaluated as the amount of variance in $\mathbf{Y}_i$ fitted by the model, for each $i \in \{1, .., n_{features}\}$. This can be computed through cross-validation, by *learning*$\widehat{\mathbf{w}}_i$ on some part of the dataset, and then estimating $\|\mathbf{Y}_i - \mathbf{X}\widehat{w}_i\|^2$ using the remainder of the dataset.

This framework is easily extended by considering

- *Grouped penalization*, where the penalization explicitly includes a prior clustering of the features, i.e. voxel-related signals, into given groups. This amounts to enforcing structured priors on the solution.
- *Combined penalizations*, i.e. a mixture of simple and group-wise penalizations, that allow some variability to fit the data in different populations of subjects, while keeping some common constraints.
- *Logistic and hinge regression*, where a non-linearity is applied to the linear model so that it yields a probability of classification in a binary classification problem.
- *Robustness to between-subject variability* to avoid the learned model overly reflecting a few outlying particular observations of the training set. Note that noise and deviating assumptions can be present in both $\mathbf{Y}$ and $\mathbf{X}$
- *Multi-task learning*: if several target variables are thought to be related, it might be useful to constrain the estimated parameter vector $\mathbf{w}$ to have a shared support across all these variables.

For instance, when one of the variables $\mathbf{Y}_i$ is not well fitted by the model, the estimation of other variables $\mathbf{Y}_j, j \neq i$ may provide constraints on the support of $\mathbf{w}_i$ and thus, improve the prediction of $\mathbf{Y}_i$.

$$\mathbf{Y} = \mathbf{X}\mathbf{w} + \epsilon, \tag{12}$$

then

$$\widehat{w} = \text{argmin}_{\mathbf{w}=(\mathbf{w}_i),i=1..n_f} \sum_{i=1}^{n_f} \|\mathbf{Y_i} - \mathbf{X}\mathbf{w_i}\|^2 + \lambda \sum_{j=1}^{n_{voxels}} \sqrt{\sum_{i=1}^{n_f} \mathbf{w_{i,j}^2}} \tag{13}$$

3.2. Multivariate decompositions

Multivariate decompositions provide a way to model complex data such as brain activation images: for instance, one might be interested in extracting an *atlas of brain regions* from a given dataset, such as regions exhibiting similar activity during a protocol, across multiple protocols, or even in the absence of protocol (during resting-state). These data can often be factorized into spatial-temporal components, and thus can be estimated through *regularized Principal Components Analysis* (PCA) algorithms, which share some common steps with regularized regression.
Let $\mathbf{X}$ be a neuroimaging dataset written as an $(n_{subjects}, n_{voxels})$ matrix, after proper centering; the model reads

$$\mathbf{X} = \mathbf{A}\mathbf{D} + \epsilon, \tag{14}$$

where $\mathbf{D}$ represents a set of n_{comp} spatial maps, hence a matrix of shape (n_{comp}, n_{voxels}), and $\mathbf{A}$ the associated subject-wise loadings. While traditional PCA and independent components analysis (ICA) are limited to reconstructing components $\mathbf{D}$ within the space spanned by the column of $\mathbf{X}$, it seems desirable to add some constraints on the rows of $\mathbf{D}$, that represent spatial maps, such as sparsity, and/or smoothness, as it makes the interpretation of these maps clearer in the context of neuroimaging. This yields the following estimation problem:

$$\text{min}_{\mathbf{D},\mathbf{A}} \|\mathbf{X} - \mathbf{A}\mathbf{D}\|^2 + \Psi(\mathbf{D}) \text{ s.t. } \|\mathbf{A}_i\| = 1 \ \forall i \in \{1..n_{features}\}, \tag{15}$$

where $(\mathbf{A}_i),\ i \in \{1..n_{features}\}$ represents the columns of $\mathbf{A}$. Ψ can be chosen such as in Eq. (2) in order to enforce smoothness and/or sparsity constraints.
The problem is not jointly convex in all the variables but each penalization given in Eq (2) yields a convex problem on $\mathbf{D}$ for $\mathbf{A}$ fixed, and conversely. This readily suggests an alternate optimization scheme, where $\mathbf{D}$ and $\mathbf{A}$ are estimated in turn, until convergence to a local optimum of the criterion. As in PCA, the extracted components can be ranked according to the amount of fitted variance. Importantly, also, estimated PCA models can be interpreted as a probabilistic model of the data, assuming a high-dimensional Gaussian distribution (probabilistic PCA).

Ultimately, the main limitations to these algorithms is the cost due to the memory requirements: holding datasets with large dimension and large number of samples (as in recent neuroimaging cohorts) leads to inefficient computation. To solve this issue, online methods are particularly attractive [24].

3.3. Covariance estimation

Another important estimation problem stems from the general issue of learning the relationship between sets of variables, in particular their covariance. Covariance learning is essential to model the dependence of these variables when they are used in a multivariate model, for instance to study potential interactions among them and with other variables. Covariance learning is necessary to model latent interactions in high-dimensional observation spaces, e.g. when considering multiple contrasts or functional connectivity data.

The difficulties are two-fold: on the one hand, there is a shortage of data to learn a good covariance model from an individual subject, and on the other hand, subject-to-subject variability poses a serious challenge to the use of multi-subject data. While the covariance structure may vary from population to population, or depending on the input data (activation versus spontaneous activity), assuming some shared structure across problems, such as their sparsity pattern, is important in order to obtain correct estimates from noisy data. Some of the most important models are:

- **Sparse Gaussian graphical models**, as they express meaningful conditional independence relationships between regions, and do improve conditioning/avoid overfit.
- **Decomposable models**, as they enjoy good computational properties and enable intuitive interpretations of the network structure. Whether they can faithfully or not represent brain networks is still an open question.
- **PCA-based regularization of covariance** which is powerful when modes of variation are more important than conditional independence relationships.

Adequate model selection procedures are necessary to achieve the right level of sparsity or regularization in covariance estimation; the natural evaluation metric here is the out-of-sample likelihood of the associated Gaussian model. Another essential remaining issue is to develop an adequate statistical framework to test differences between covariance models in different populations. To do so, we consider different means of parametrizing covariance distributions and how these parametrizations impact the test of statistical differences across individuals.

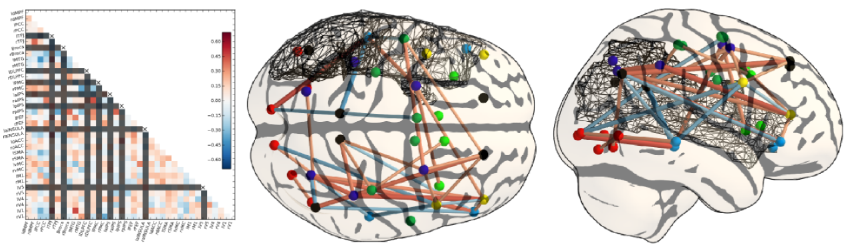


Figure 2. Example of functional connectivity analysis: The correlation matrix describing brain functional connectivity in a post-stroke patient (lesion volume outlined as a mesh) is compared to a group of control subjects. Some edges of the graphical model show a significant difference, but the statistical detection of the difference requires a sophisticated statistical framework for the comparison of graphical models.

4. Application Domains

4.1. Cognitive neuroscience

4.1.1. Macroscopic Functional cartography with functional Magnetic Resonance Imaging (fMRI)

The brain as a highly structured organ, with both functional specialization and a complex network organization. While most of the knowledge historically comes from lesion studies and animal electophysiological recordings, the development of non-invasive imaging modalities, such as fMRI, has made it possible to study routinely high-level cognition in humans since the early 90's. This has opened major questions on the interplay between mind and brain , such as: How is the function of cortical territories constrained by anatomy (connectivity) ? How to assess the specificity of brain regions ? How can one characterize reliably inter-subject differences ?

4.1.2. Analysis of brain Connectivity

Functional connectivity is defined as the interaction structure that underlies brain function. Since the beginning of fMRI, it has been observed that remote regions sustain high correlation in their spontaneous activity, i.e. in the absence of a driving task. This means that the signals observed during resting-state define a signature of the connectivity of brain regions. The main interest of resting-state fMRI is that it provides easy-to-acquire functional markers that have recently been proved to be very powerful for population studies.

4.1.3. Modeling of brain processes (MEG)

While fMRI has been very useful in defining the function of regions at the mm scale, Magneto-encephalography (MEG) provides the other piece of the puzzle, namely temporal dynamics of brain activity, at the ms scale. MEG is also non-invasive. It makes it possible to keep track of precise schedule of mental operations and their interactions. It also opens the way toward a study of the rhythmic activity of the brain. On the other hand, the localization of brain activity with MEG entails the solution of a hard inverse problem.

4.1.4. Current challenges in human neuroimaging (acquisition+analysis)

Human neuroimaging targets two major goals: *i)* the study of neural responses involved in sensory, motor or cognitive functions, in relation to models from cognitive psychology, i.e. the identification of neurophysiological and neuroanatomical correlates of cognition; *ii)* the identification of markers in brain structure and function of neurological or psychiatric diseases. Both goals have to deal with a tension between

- the search for higher spatial [0] resolution to increase **spatial specificity** of brain signals, and clarify the nature (function and structure) of brain regions. This motivates efforts for high-field imaging and more efficient acquisitions, such as compressed sensing schemes, as well as better source localization methods from M/EEG data.
- the importance of inferring brain features with **population-level** validity, hence, contaminated with high variability within observed cohorts, which blurs the information at the population level and ultimately limits the spatial resolution of these observations.

Importantly, the signal-to-noise ratio (SNR) of the data remains limited due to both resolution improvements [0] and between-subject variability. Altogether, these factors have led to realize that results of neuroimaging studies were **statistically weak**, i.e. plagued with low power and leading to unreliable inference [70], and particularly so due to the typically number of subjects included in brain imaging studies (20 to 30, this number tends to increase [71]): this is at the core of the *neuroimaging reproducibility crisis*. This crisis is deeply related to a second issue, namely that only few neuroimaging datasets are publicly available, making it impossible to re-assess a posteriori the information conveyed by the data. Fortunately, the situation improves, lead by projects such as NeuroVault or OpenfMRI. A framework for integrating such datasets is however still missing.

[0] and to some extent, temporal, but for the sake of simplicity we focus here on spatial aspects.

[0] The SNR of the acquired signal is proportional to the voxel size, hence an improvement by a factor of 2 in image resolution along each dimension is payed by a factor of 8 in terms of SNR.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

- Pierre Ablin got a best student paper award at the LVA-ICA conference for his paper [34].
- First PhD prize from STIC doctoral school for Tom Dupré la Tour.

6. New Software and Platforms

6.1. Mayavi

FUNCTIONAL DESCRIPTION: Mayavi is the most used scientific 3D visualization Python software. Mayavi can be used as a visualization tool, through interactive command line or as a library. It is distributed under Linux through Ubuntu, Debian, Fedora and Mandriva, as well as in PythonXY and EPD Python scientific distributions. Mayavi is used by several software platforms, such as PDE solvers (fipy, sfepy), molecule visualization tools and brain connectivity analysis tools (connectomeViewer).

- Contact: Gaël Varoquaux
- URL: http://mayavi.sourceforge.net/

6.2. MedInria

KEYWORDS: Visualization - DWI - Health - Segmentation - Medical imaging
SCIENTIFIC DESCRIPTION: medInria aims at creating an easily extensible platform for the distribution of research algorithms developed at Inria for medical image processing. This project has been funded by the D2T (ADT MedInria-NT) in 2010, renewed in 2012. A fast-track ADT was awarded in 2017 to transition the software core to more recent dependencies and study the possibility of a consortium creation.The Visages team leads this Inria national project and participates in the development of the common core architecture and features of the software as well as in the development of specific plugins for the team's algorithm.
FUNCTIONAL DESCRIPTION: MedInria is a free software platform dedicated to medical data visualization and processing.

- Participants: Maxime Sermesant, Olivier Commowick and Théodore Papadopoulo
- Partners: HARVARD Medical School - IHU - LIRYC - NIH
- Contact: Olivier Commowick
- URL: http://med.inria.fr

6.3. Nilearn

NeuroImaging with scikit learn
KEYWORDS: Health - Neuroimaging - Medical imaging
FUNCTIONAL DESCRIPTION: NiLearn is the neuroimaging library that adapts the concepts and tools of scikit-learn to neuroimaging problems. As a pure Python library, it depends on scikit-learn and nibabel, the main Python library for neuroimaging I/O. It is an open-source project, available under BSD license. The two key components of NiLearn are i) the analysis of functional connectivity (spatial decompositions and covariance learning) and ii) the most common tools for multivariate pattern analysis. A great deal of efforts has been put on the efficiency of the procedures both in terms of memory cost and computation time.

- Participants: Alexandre Abraham, Alexandre Gramfort, Bertrand Thirion, Elvis Dohmatob, Fabian Pedregosa Izquierdo, Gaël Varoquaux, Loïc Estève, Michael Eickenberg and Virgile Fritsch
- Contact: Bertrand Thirion
- URL: http://nilearn.github.io/

6.4. PyHRF

KEYWORDS: Medical imaging - Health - Brain - IRM - Neurosciences - Statistic analysis - FMRI
SCIENTIFIC DESCRIPTION: Functional Magnetic Resonance Imaging (fMRI) is a neuroimaging technique that allows the non-invasive study of brain function. It is based on the hemodynamic variations induced by changes in cerebral synaptic activity following sensory or cognitive stimulation. The measured signal depends on the variation of blood oxygenation level (BOLD signal) which is related to brain activity: a decrease in deoxyhemoglobin concentration induces an increase in BOLD signal. The BOLD signal is delayed with respect to changes in synaptic activity, which can be modeled as a convolution with the Hemodynamic Response Function (HRF) whose exact form is unknown and fluctuates with various parameters such as age, brain region or physiological conditions. In this work we propose to analyze fMRI data using a Joint Detection-Estimation (JDE) approach. It jointly detects cortical activation and estimates the HRF. In contrast to existing tools, PyHRF estimates the HRF instead of considering it as a given constant in the entire brain.
FUNCTIONAL DESCRIPTION: As part of fMRI data analysis, PyHRF provides a set of tools for addressing the two main issues involved in intra-subject fMRI data analysis : (i) the localization of cerebral regions that elicit evoked activity and (ii) the estimation of the activation dynamics also referenced to as the recovery of the Hemodynamic Response Function (HRF). To tackle these two problems, PyHRF implements the Joint Detection-Estimation framework (JDE) which recovers parcel-level HRFs and embeds an adaptive spatio-temporal regularization scheme of activation maps.
NEWS OF THE YEAR: The framework to perform software tests has been further developed. Some unitary tests have been set.

- Participants: Aina Frau Pascual, Christine Bakhous, Florence Forbes, Jaime Eduardo Arias Almeida, Laurent Risser, Lotfi Chaari, Philippe Ciuciu, Solveig Badillo, Thomas Perret and Thomas Vincent
- Partners: CEA - NeuroSpin
- Contact: Florence Forbes
- Publications: Frontiers in Neuroinformatics Flexible multivariate hemodynamics fMRI data analyses and simulations with PyHRF - Fast joint detection-estimation of evoked brain activity in event-related fMRI using a variational approach - A Bayesian Non-Parametric Hidden Markov Random Model for Hemodynamic Brain Parcellation
- URL: http://pyhrf.org

6.5. Scikit-learn

KEYWORDS: Regession - Clustering - Learning - Classification - Medical imaging
SCIENTIFIC DESCRIPTION: Scikit-learn is a Python module integrating classic machine learning algorithms in the tightly-knit scientific Python world. It aims to provide simple and efficient solutions to learning problems, accessible to everybody and reusable in various contexts: machine-learning as a versatile tool for science and engineering.
FUNCTIONAL DESCRIPTION: Scikit-learn can be used as a middleware for prediction tasks. For example, many web startups adapt Scikitlearn to predict buying behavior of users, provide product recommendations, detect trends or abusive behavior (fraud, spam). Scikit-learn is used to extract the structure of complex data (text, images) and classify such data with techniques relevant to the state of the art.

Easy to use, efficient and accessible to non datascience experts, Scikit-learn is an increasingly popular machine learning library in Python. In a data exploration step, the user can enter a few lines on an interactive (but non-graphical) interface and immediately sees the results of his request. Scikitlearn is a prediction engine . Scikit-learn is developed in open source, and available under the BSD license.

- Participants: Alexandre Gramfort, Bertrand Thirion, Fabian Pedregosa Izquierdo, Gaël Varoquaux, Loïc Estève, Michael Eickenberg and Olivier Grisel
- Partners: CEA - Logilab - Nuxeo - Saint Gobain - Tinyclues - Telecom Paris
- Contact: Olivier Grisel
- URL: http://scikit-learn.org

6.6. MODL

Massive Online Dictionary Learning
KEYWORDS: Pattern discovery - Machine learning
FUNCTIONAL DESCRIPTION: Matrix factorization library, usable on very large datasets, with optional sparse and positive factors.

- Participants: Arthur Mensch, Gaël Varoquaux, Bertrand Thirion and Julien Mairal
- Contact: Arthur Mensch
- Publications: Subsampled online matrix factorization with convergence guarantees - Stochastic Subsampling for Factorizing Huge Matrices
- URL: http://github.com/arthurmensch/modl

6.7. MNE

MNE-Python
KEYWORDS: Neurosciences - EEG - MEG - Signal processing - Machine learning
FUNCTIONAL DESCRIPTION: Open-source Python software for exploring, visualizing, and analyzing human neurophysiological data: MEG, EEG, sEEG, ECoG, and more.

RELEASE FUNCTIONAL DESCRIPTION: http://martinos.org/mne/stable/what_new.html

- Partners: HARVARD Medical School - New York University - University of Washington - CEA - Aalto university - Telecom Paris - Boston University - UC Berkeley
- Contact: Alexandre Gramfort
- URL: http://martinos.org/mne/

7. New Results

7.1. Reducing the number of samples in spatiotemporal dMRI acquisition design

Acquisition time is a major limitation in recovering brain white matter microstructure with diffusion magnetic resonance imaging. The aim of this work is to bridge the gap between growing demands on spatio-temporal resolution of diffusion signal and the real-world time limitations. We introduce an acquisition scheme that reduces the number of samples under adjustable quality loss. Finding a sampling scheme that maximizes signal quality and satisfies given time constraints is NP-hard. Therefore, a heuristic method based on genetic algorithm is proposed in order to find sub-optimal solutions in acceptable time. The analyzed diffusion signal representation is defined in the qτ space, so that it captures both spacial and temporal phenomena. The experiments on synthetic data and in vivo diffusion images of the C57Bl6 wild-type mouse corpus callosum reveal the superiority of the proposed approach over random sampling and even distribution in the qτ space. The use of genetic algorithm allows to find acquisition parameters that guarantee high signal reconstruction accuracy under given time constraints. In practice, the proposed approach helps to accelerate the acquisition for the use of q-dMRI signal representation.

More information can be found in [12]

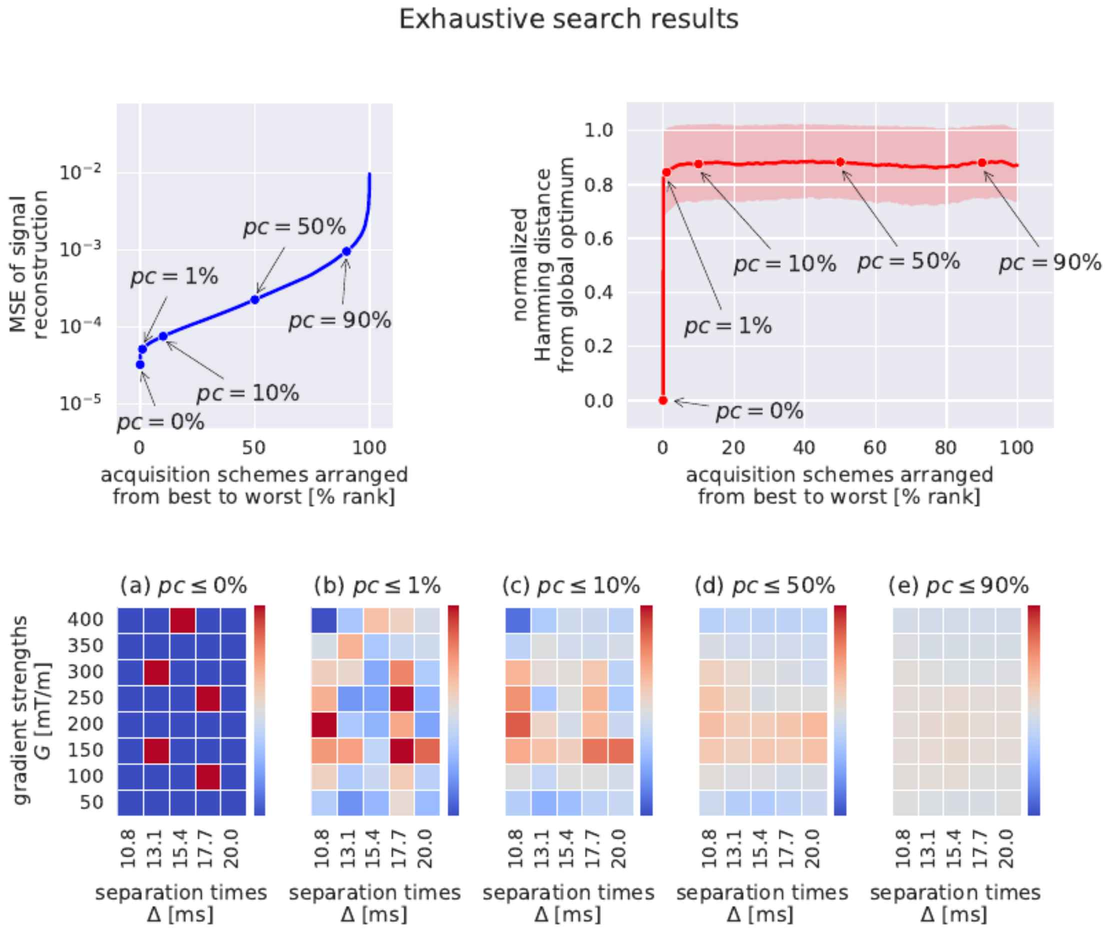


Figure 3. Exhaustive search results of the optimization by shells for the in silico experiment with nmax = 100. The plots at the top present all the 658,008 feasible acquisition schemes arranged from best to worst, illustrating the mean squared errors (MSEs) of signal reconstruction (top-left plot) and the normalized Hamming distances from the global optimum $\pm$ 1 standard deviation (top-right). In order to visualize the analyzed (G, Δ) parameter space, the percentiles pc = 0%, 1%, 10%, 50%, 90% are annotated on both plots, showing respectively the global optimum, the top 1% solutions, the top 10% solutions, etc. The corresponding cumulative averages of acquisition schemes are depicted in the heat maps at the bottom. The colors reflect the likelihood of a given (G, Δ) pair in the scheme. The heat maps for pc $\leq$ 0% and pc $\leq$ 1% represent, respectively, the global optimum and its proximity. The interval between pc = 10% and pc = 90% contains a huge spectrum of schemes with similar MSEs and almost equally large distances from the global optimum.

7.2. Robust EEG-based cross-site and cross-protocol classification of states of consciousness

Determining the state-of-consciousness in patients with disorders-of-consciousness (DOC) is a challenging practical and theoretical problem. Recent findings suggest that multiple markers of brain activity extracted from the electroencephalogram (EEG) may index the state of consciousness in the human brain. Furthermore, machine learning has been found to optimize their capacity to discriminate different states of consciousness in clinical practice. However, it is unknown how dependable these EEG-markers are in the face of signal variability due to different EEG-configurations, EEG-protocols and subpopulations from different centers encountered in practice. In our recent paper [11] we addressed the following questions: What is the impact of the EEG configuration (selection of sensors, duration of EEG used)? Do models based on current EEG-markers achieve prospective generalization on independent data from other EEG protocols and other hospitals? Are single markers sufficiently powerful and when does multivariate classification provide the clearest advantage? For summary of methods and approahc see Figure4. Our results highlight the effectiveness of classical well-studied EEG-signatures such as alpha [8-12Hz] and theta [5-7Hz] frequency band oscillations for detecting consciousness when combined with machine learning. While univariate predictive models achieved good performance, multivariate models showed better generalization capacity and increased robustness to different types of noise while mitigating the impact of the EEG-configuration. Our findings suggest that pooling data over multiple centers for predictive modeling of DOC is a concrete possibility and can become a promising alley for the field of cognitive neurology.

7.3. A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series

Sleep stage classification constitutes an important preliminary exam in the diagnosis of sleep disorders. It is traditionally performed by a sleep expert who assigns to each 30 s of signal a sleep stage, based on the visual inspection of signals such as electroencephalograms (EEG), electrooculograms (EOG), electrocardiograms (ECG) and electromyograms (EMG). We introduce here the first deep learning approach for sleep stage classification that learns end-to-end without computing spectrograms or extracting hand-crafted features, that exploits all multivariate and multimodal Polysomnography (PSG) signals (EEG, EMG and EOG), and that can exploit the temporal context of each 30 s window of data. For each modality the first layer learns linear spatial filters that exploit the array of sensors to increase the signal-to-noise ratio, and the last layer feeds the learnt representation to a softmax classifier. Our model is compared to alternative automatic approaches based on convolutional networks or decisions trees. Results obtained on 61 publicly available PSG records with up to 20 EEG channels demonstrate that our network architecture yields state-of-the-art performance. Our study reveals a number of insights on the spatio-temporal distribution of the signal of interest: a good trade-off for optimal classification performance measured with balanced accuracy is to use 6 EEG with 2 EOG (left and right) and 3 EMG chin channels. Also exploiting one minute of data before and after each data segment offers the strongest improvement when a limited number of channels is available. As sleep experts, our system exploits the multivariate and multimodal nature of PSG signals in order to deliver state-of-the-art classification performance with a small computational cost.

More information can be found in [8].

7.4. Individual Brain Charting, a high-resolution fMRI dataset for cognitive mapping

Functional Magnetic Resonance Imaging (fMRI) has furthered brain mapping on perceptual, motor, as well as higher-level cognitive functions. However, to date, no data collection has systematically addressed the functional mapping of cognitive mechanisms at a fine spatial scale. The Individual Brain Charting (IBC) project stands for a high-resolution multi-task fMRI dataset that intends to provide the objective basis toward a comprehensive functional atlas of the human brain. The data refer to a cohort of 12 participants performing

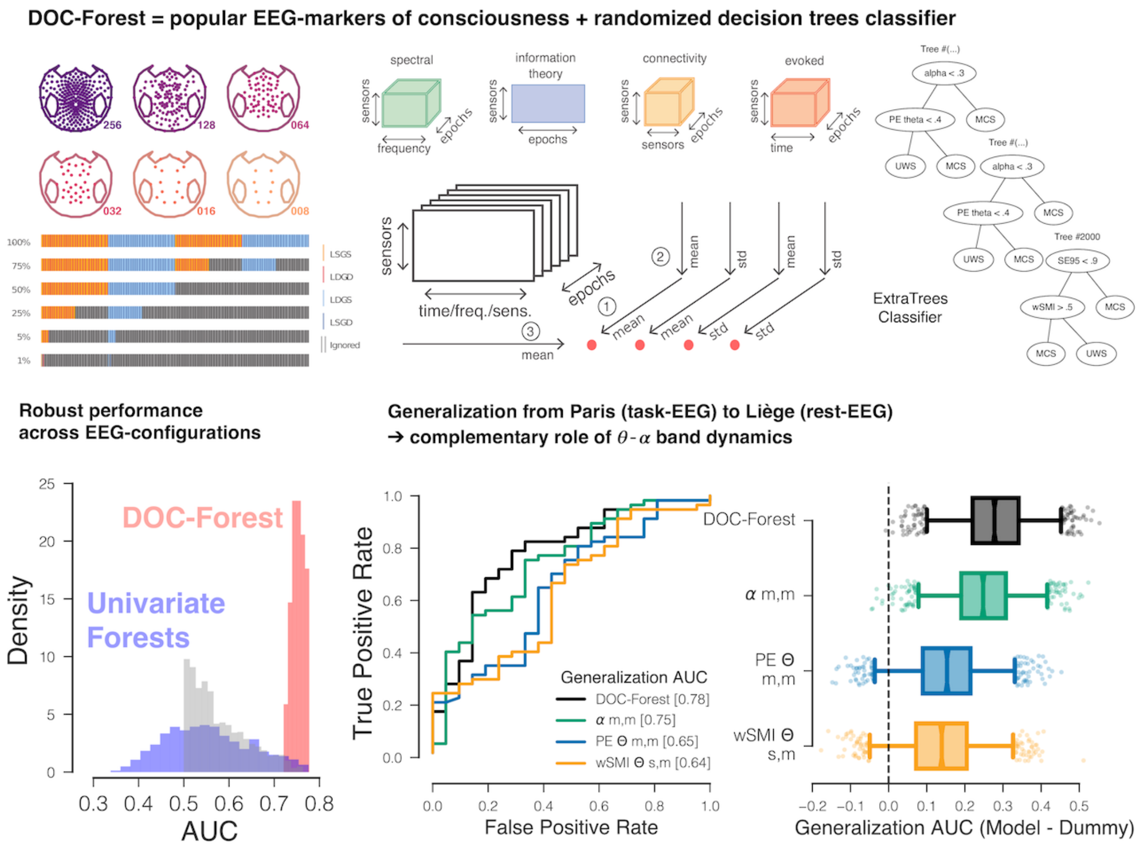


Figure 4. We probed the robustness and validity of EEG-markers of consciousness. Using the robust Extra-Trees algorithm (Geurts, Ernst, & Wehenkel, 2006) we developed a classifier trained to differentiate UWS from MCS patients. This classifier (named "DOC-forest") was trained and tested using 28 potential EEG-markers of consciousness (112 features) from 249 patients recorded at the Paris Pitié-Salpêtrière and 78 patients from the University Hospital of Liège. We used the MNE-Python software for EEG processing and the scikit-learn package for machine learning. Our results show that optimally combining multiple EEG-markers of states of consciousness using machine learning enables robust generalization across EEG-configurations, EEG-protocols and sites. Our recipe for extracting biomarkers is available on Github: https://nice-tools.github.io/nice. For a neuroscientific discussion of our work see the accompanying commentary article by Sokoliuk and Cruse (https://doi.org/10.1093/brain/awy267).

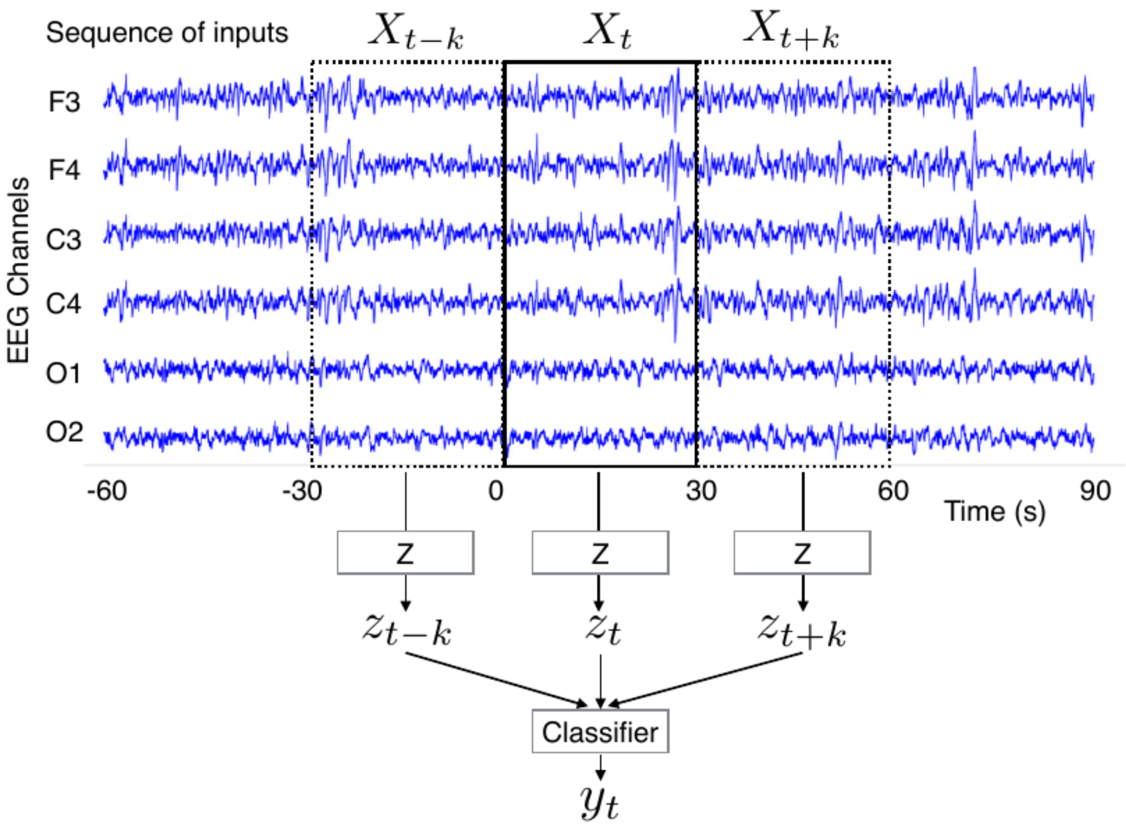


Figure 5. Time distributed architecture to process a sequence of inputs $S_t^k = \{X_{t-k}, \cdots, X_t, \cdots, X_{t+k}\}$ with k = 1. X_k stands for the multivariate input data over 30 s that is fed into the feature extractor Z. Features are extracted from consecutive 30 s samples: X_{t-k}, ..., X_t, ..., X_{t+k}. Then the obtained features are aggregated $[z_{t-k}, \cdots, z_t, \cdots, z_{t+k}]$. The resulting aggregation of features is finally fed into a classifier to predict the label y_t associated with the sample X_t.

many different tasks. The large amount of task-fMRI data on the same subjects yields a precise mapping of the underlying functions, free from both inter-subject and inter-site variability. The present article gives a detailed description of the first release of the IBC dataset. It comprises a dozen of tasks, addressing both low- and high-level cognitive functions. This openly available dataset is thus intended to become a reference for cognitive brain mapping.

More information can be found in [25]

7.5. Atlases of cognition with large-scale brain mapping

To map the neural substrate of mental function, cognitive neuroimaging relies on controlled psychological manipulations that engage brain systems associated with specific cognitive processes. In order to build comprehensive atlases of cognitive function in the brain, it must assemble maps for many different cognitive processes, which often evoke overlapping patterns of activation. Such data aggregation faces contrasting goals: on the one hand finding correspondences across vastly different cognitive experiments, while on the other hand precisely describing the function of any given brain region. Here we introduce a new analysis framework that tackles these difficulties and thereby enables the generation of brain atlases for cognitive function. The approach leverages ontologies of cognitive concepts and multi-label brain decoding to map the neural substrate of these concepts. We demonstrate the approach by building an atlas of functional brain organization based on 30 diverse functional neuroimaging studies, totaling 196 different experimental conditions. Unlike conventional brain mapping, this functional atlas supports robust reverse inference: predicting the mental processes from brain activity in the regions delineated by the atlas. To establish that this reverse inference is indeed governed by the corresponding concepts, and not idiosyncrasies of experimental designs, we show that it can accurately decode the cognitive concepts recruited in new tasks. These results demonstrate that aggregating independent task-fMRI studies can provide a more precise global atlas of selective associations between brain and cognition.

More information can be found in [28].

7.6. Celer: a Fast Solver for the Lasso with Dual Extrapolation

Convex sparsity-inducing regularizations are ubiquitous in high-dimensional machine learning, but solving the resulting optimization problems can be slow. To accelerate solvers, state-of-the-art approaches consist in reducing the size of the optimization problem at hand. In the context of regression, this can be achieved either by discarding irrelevant features (screening techniques) or by prioritizing features likely to be included in the support of the solution (working set techniques). Convex duality comes into play at several steps in these techniques. Here, we propose an extrapolation technique starting from a sequence of iterates in the dual that leads to the construction of improved dual points. This enables a tighter control of optimality as used in stopping criterion, as well as better screening performance of Gap Safe rules. Finally, we propose a working set strategy based on an aggressive use of Gap Safe screening rules. Thanks to our new dual point construction, we show significant computational speedups on multiple real-world problems compared to alternative state-of-the-art coordinate descent solvers.

More information can be found in [54]. Code can be found at https://mathurinm.github.io/celer/.

7.7. Multivariate Convolutional Sparse Coding for Electromagnetic Brain Signals

Frequency-specific patterns of neural activity are traditionally interpreted as sustained rhythmic oscillations, and related to cognitive mechanisms such as attention, high level visual processing or motor control. While alpha waves (8–12 Hz) are known to closely resemble short sinusoids, and thus are revealed by Fourier analysis or wavelet transforms, there is an evolving debate that electromagnetic neural signals are composed of more complex waveforms that cannot be analyzed by linear filters and traditional signal representations. In this work, we propose to learn dedicated representations of such recordings using a multivariate convolutional sparse

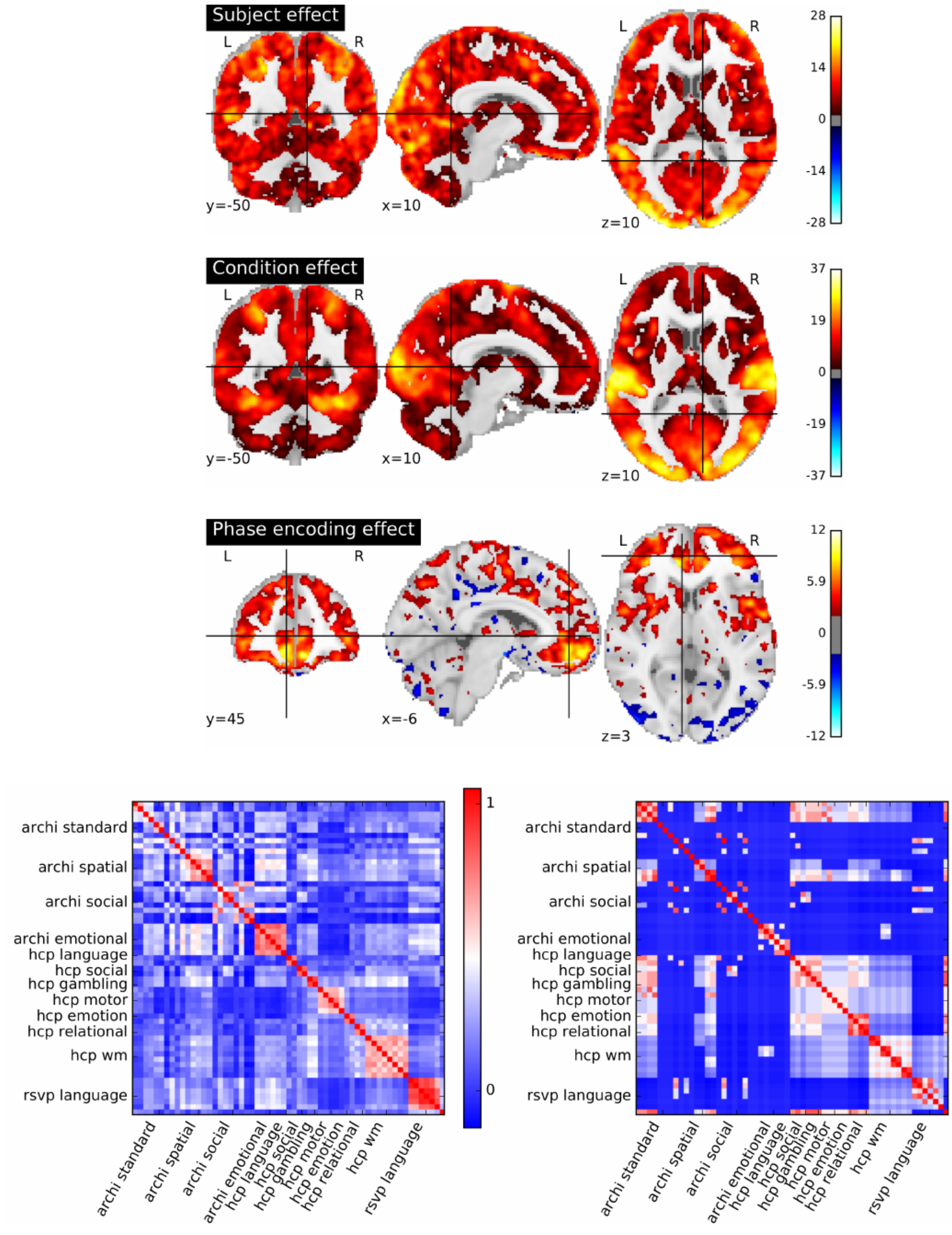


Figure 6. Overview of information conveyed by activation maps resulting from a first-level analysis. (top) Global effects of experimental subject condition, and phase-encoding direction. A per-voxel ANOVA breaks the variance of the set of brain maps into subject, experimental condition, and phase-encoding direction values. All maps are given in z-scale and thresholded at an FDR level of 0.05. (Bottom) Focusing on condition effect, the similarity between condition-related maps, averaged across subjects (left) is clearly related to the dissimilarity of the conditions, when these are characterized in terms of the Cognitive Atlas (right).

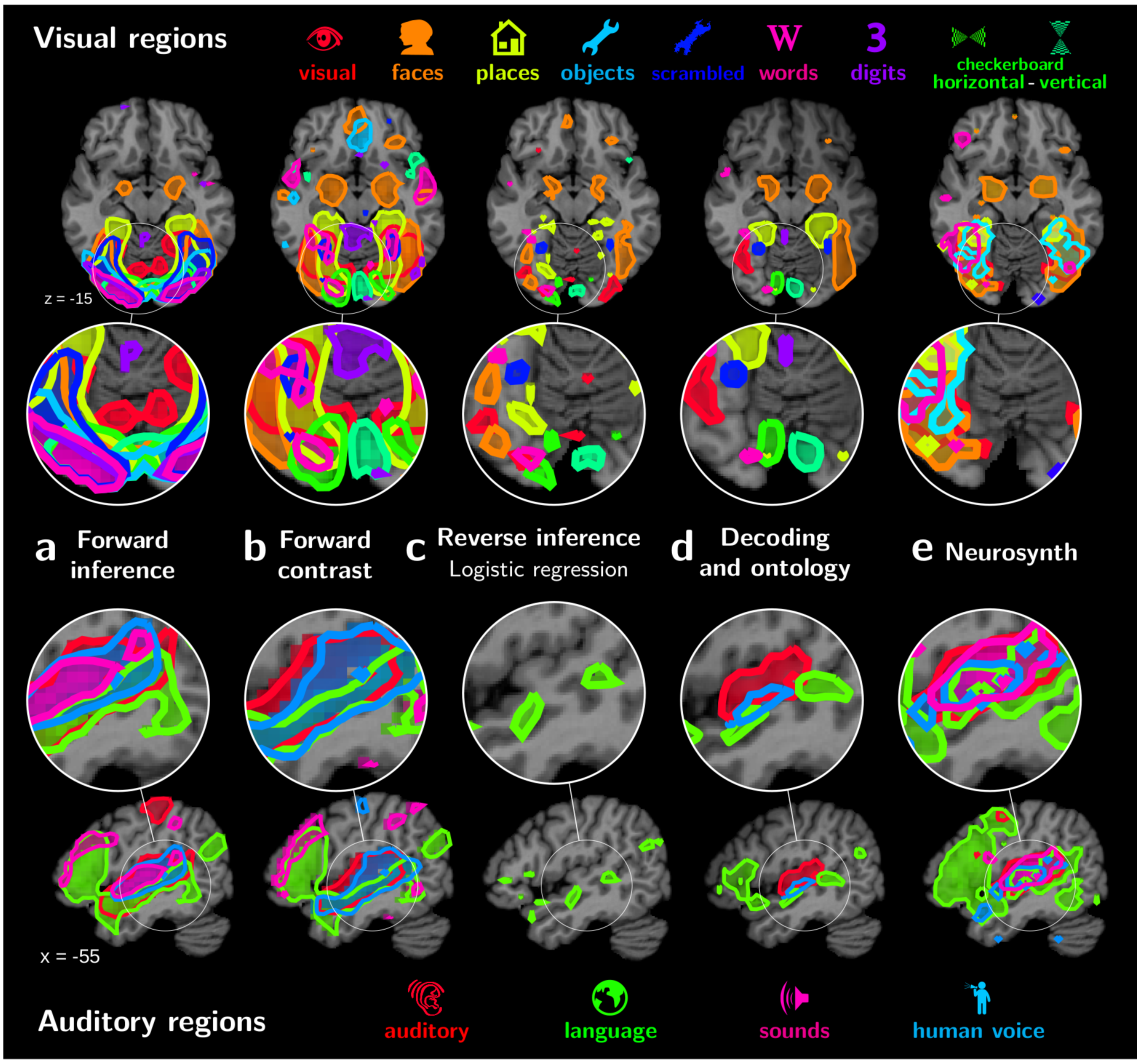


Figure 7. Different functional atlases – Regions outlined using different functional mapping approaches, from left to right: a. forward term mapping; b. forward inference with ontology contrasts (standard analysis); c. reverse inference with logistic regression; d. NeuroSynth reverse inference; and e. our approach, mapping with decoding and an ontology. The top part shows visual regions, and the lower one auditory regions in the left hemisphere. Forward term mapping outlines overlapping regions, as brain responses capture side effects such as the stimulus modality: for visual and auditory regions every cognitive term is represented in the corresponding primary cortex. Forward mapping using contrasts removes the overlap in primary regions, but a large overlap persists in mid-level regions, as control conditions are not well matched across studies. Standard reverse inference, specific to a term, creates overly sparse regions though with little overlap. Reverse inference with Neurosynth also displays large overlap in mid-level regions. Finally, ontology-based decoding maps recover known functional areas the visual and auditory cortices.

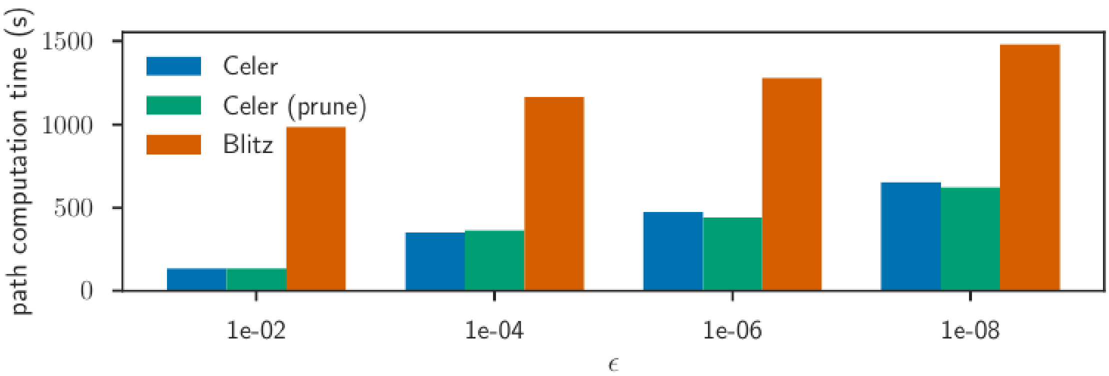


Figure 8. Times to solve the Lasso path to precision ϵ for 100 values of λ, from λ_{max} to $\lambda_{max}/100$, on the Finance data. CELER outperforms BLITZ . Both safe and prune versions behave similarly.

coding (CSC) algorithm. Applied to electroencephalography (EEG) or magnetoencephalography (MEG) data, this method is able to learn not only prototypical temporal waveforms, but also associated spatial patterns so their origin can be localized in the brain. Our algorithm is based on alternated minimization and a greedy coordinate descent solver that leads to state-of-the-art running time on long time series. To demonstrate the implications of this method, we apply it to MEG data and show that it is able to recover biological artifacts. More remarkably, our approach also reveals the presence of non-sinusoidal mu-shaped patterns, along with their topographic maps related to the somatosensory cortex.

More information can be found in [52]. Code can be found at https://alphacsc.github.io/.

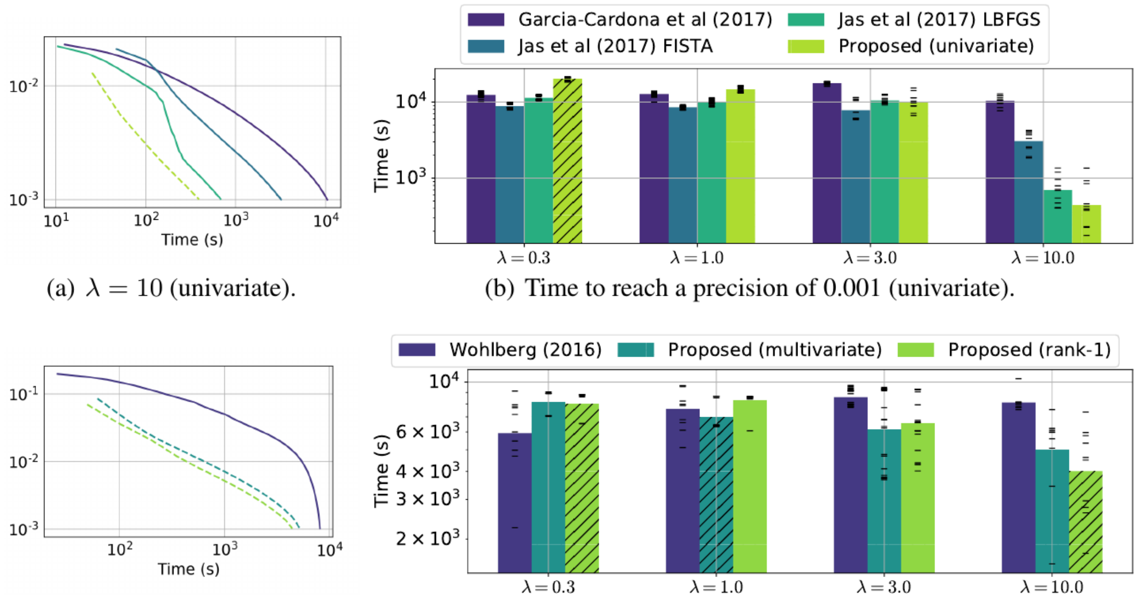
7.8. Stochastic Subsampling for Factorizing Huge Matrices

We present a matrix-factorization algorithm that scales to input matrices with both huge number of rows and columns. Learned factors may be sparse or dense and/or non-negative, which makes our algorithm suitable for dictionary learning, sparse component analysis, and non-negative matrix factorization. Our algorithm streams matrix columns while subsampling them to iteratively learn the matrix factors. At each iteration, the row dimension of a new sample is reduced by subsampling, resulting in lower time complexity compared to a simple streaming algorithm. Our method comes with convergence guarantees to reach a stationary point of the matrix-factorization problem. We demonstrate its efficiency on massive functional Magnetic Resonance Imaging data (2 TB), and on patches extracted from hyperspectral images (103 GB). For both problems, which involve different penalties on rows and columns, we obtain significant speed-ups compared to state-of-the-art algorithms.

More information can be found in [24].

7.9. Text to brain: predicting the spatial distribution of neuroimaging observations from text reports

Despite the digital nature of magnetic resonance imaging, the resulting observations are most frequently reported and stored in text documents. There is a trove of information untapped in medical health records, case reports, and medical publications. In this paper, we propose to mine brain medical publications to learn the spatial distribution associated with anatomical terms. The problem is formulated in terms of minimization of a risk on distributions which leads to a least-deviation cost function. An efficient algorithm in the dual then learns the mapping from documents to brain structures. Empirical results using coordinates extracted from the



(a) $\lambda = 10$ (univariate).

(b) Time to reach a precision of 0.001 (univariate).

(c) $\lambda = 10$ (multivariate).

(d) Time to reach a precision of 0.001 (multivariate).

Figure 9. Comparison of state-of-the-art univariate (a, b) and multivariate (c, d) methods with our approach. (a) Convergence plot with the objective function relative to the obtained minimum, as a function of computational time. (b) Time taken to reach a relative precision of 10^{-3}, for different regularization parameters λ. (c, d) Same as (a, b) in the multivariate setting P=5.

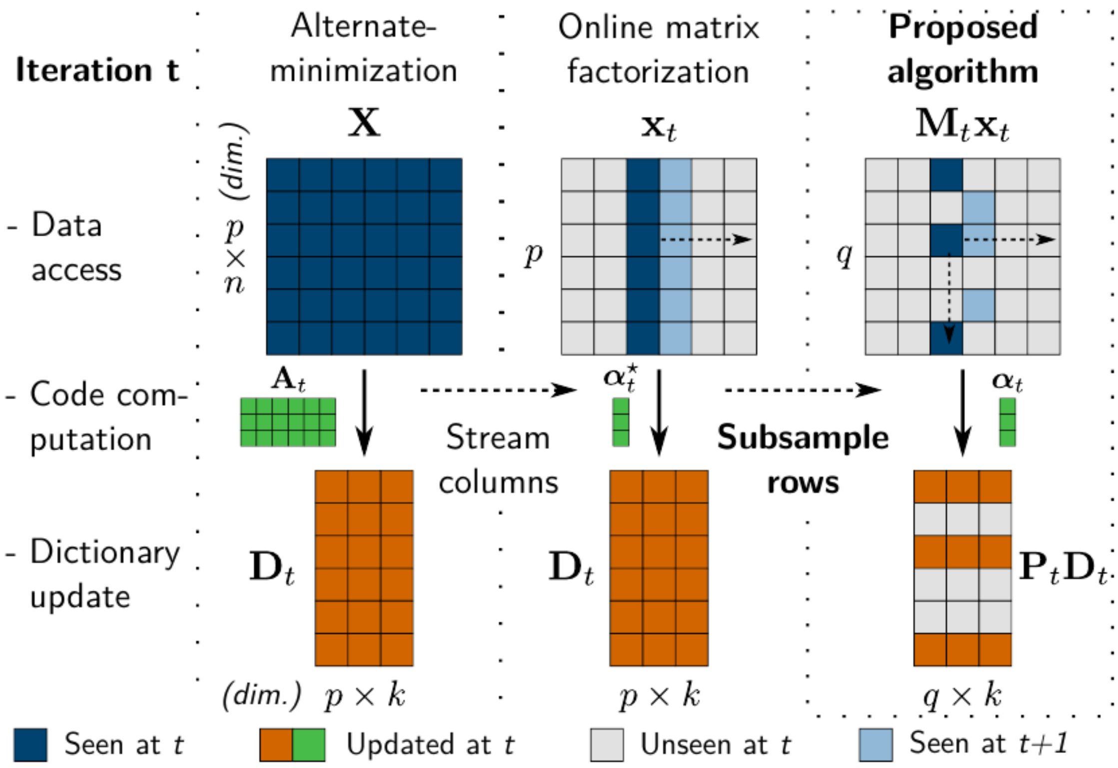


Figure 10. Stochastic subsampling further improves online matrix factorizationhandle datasets with large number of columns and rows. X is the input $p \times n$ matrix, D_t and A_t are respectively the dictionary and code at time t.

brain-imaging literature show that i) models must adapt to semantic variation in the terms used to describe a given anatomical structure, ii) voxel-wise parameterization leads to higher likelihood of locations reported in unseen documents, iii) least-deviation cost outperforms least-square. As a proof of concept for our method, we use our model of spatial distributions to predict the distribution of specific neurological conditions from text-only reports.

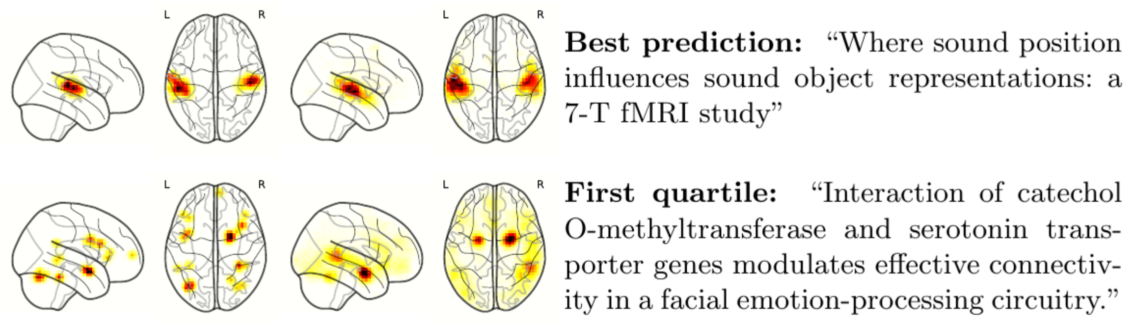


Figure 11. True probability density function (estimated with kernel density estimator) and the prediction for the articles which obtained respectively the best and the first- quartile scores.

More information can be found in [37].

7.10. Similarity encoding for learning with dirty categorical variables

For statistical learning, categorical variables in a table are usually considered as discrete entities and encoded separately to feature vectors, e.g., with one-hot encoding. "Dirty" non-curated data gives rise to categorical variables with a very high cardinality but redundancy: several categories reflect the same entity. In databases, this issue is typically solved with a deduplication step. We show that a simple approach that exposes the redundancy to the learning algorithm brings significant gains. We study a generalization of one-hot encoding, similarity encoding, that builds feature vectors from similarities across categories. We perform a thorough empirical validation on non-curated tables, a problem seldom studied in machine learning. Results on seven real-world datasets show that similarity encoding brings significant gains in prediction in comparison with known encoding methods for categories or strings, notably one-hot encoding and bag of character n-grams. We draw practical recommendations for encoding dirty categories: 3-gram similarity appears to be a good choice to capture morphological resemblance. For very high-cardinality, dimensionality reduction significantly reduces the computational cost with little loss in performance: random projections or choosing a subset of prototype categories still outperforms classic encoding approaches.

More information can be found in [7].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

In 2018, a CIFRE PhD thesis was launched with the Canadian company Interaxon https://choosemuse.com. This contract supports the PhD thesis of Hubert Banville.

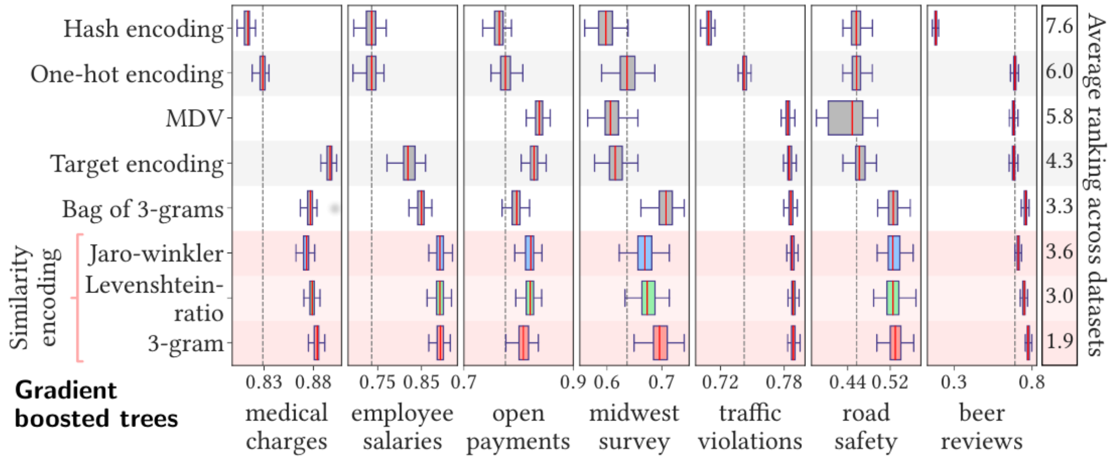


Figure 12. Performance of different encoding methods in a gradient boosting classification task. Each box-plot summarizes the prediction scores of 100 random splits (with 80% of the samples for training and 20% for testing). For all datasets, the prediction score is upper bounded by 1 (a higher score means a better prediction). The right side of the figure indicates the average ranking across datasets for each method. The vertical dashed line indicates the median value of the one-hot encoding method.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. Inserm-Inria project

This project is funded by the joint Inserm and Inria program ‘médecine numérique’ and is conducted in collaborations with our clinical partners from the Lariboisière hospital, Inserm uni U942 BioCANVAS (Biomarkers in Cardio-Neuro-VAScular diseases). It supports the PhD thesis of David Sabbagh.

Participants:

- Denis Engemann [coordinator, co-advisor]
- Alexandre Gramfort [thesis director, co-advisor]
- Etienne Gayat [clinical collaborator, co-advisor]
- Fabrice Vallée [clinical collaborator]
- David Sabbagh [PhD Student]

Post-operative delirium (POD) is a potential complication of anesthesia during surgery. It is often associated with adverse outcomes and is aggravated by aging. In elderly patients, post-operative complications have been estimated to incur tens of million US dollars of costs each year in the United States by prolonging hospitalization and potentially affecting health prognosis. Recent studies suggest that POD can already be prevented by improving electrophysiological monitoring of anesthesia depth and individual dosage of anesthetic agents. Doing so probably minimizes the time patients spend in a coma-like state that manifests itself in isoelectric burst suppression, an electroencephalogram (EEG) pattern characterized by alternation between quiescence and high-amplitude bursts, and causally linked to POD. However, such an enterprise, currently, depends on the trained clinical electrophysiologist and guidance by commercially provided EEG indices of states of consciousness. One such metric is the bispectral index (BIS), which, like other related metrics, does not explicitly take into account baseline changes related to normative aging and may therefore be biased when used naively.

While electrophysiological signatures of aging (e.g. drop in Alpha and Gamma band power), states of consciousness (e.g. drop in Theta band long-range connectivity) and drug response (e.g. anteriorization of alpha band power in propofol anesthesia) have been separately investigated in the past years, their common denominators are not known. It is therefore difficult to detect individual risk, choose the optimal dosage, and automate anesthesia monitoring readily for any patient in any hospital.

The goal of this research project is to build statistical models that enable prediction of burst suppression and subsequent POD by exploiting diverse EEG-signatures of states of consciousness in the context of aging. We approach this challenge by recasting it as a problem of learning brain-age from the point of view of electrophysiology of consciousness.

9.1.2. CoSmic project

Participants: Philippe Ciuciu [Correspondant], Carole Lazarus, Loubna El Gueddari.

This project is funded by CEA DRF-Impulsion.

This is a collaborative project with Jean-Luc Stark, (CEA) funded by the DRF-impulsion CEA program.

Compressed Sensing is a recent theory in maths that allows the perfect recovery of signals or images from compressive acquisition scenarios. This approach has been popularized in MRI over the last decade as well as in astrophysics (noticeably in radio-astronomy). So far, both of these fields have developed skills in CS separately. The aim of the COSMIC project is to foster collaborations between CEA experts in MRI (Parietal team within NeuroSpin) and in astrophysics (CosmoStat lab within the Astrophysics Department). These interactions will allow us to share different expertise in order to improve image quality, either in MRI or in radio-astronomy (thanks to the interferometry principle). In this field, given the data delivered by radio-telescopes, the goal consists in extracting high temporal resolution information in order to study fast transient events.

9.1.3. Metacog

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux, Jérome Dockès.

This project is funded by Digiteo.

This is a Digicosme project (2016-2019) and a collaboration with Fabian Suchanek (Telecom Paritech).

Understanding how cognition emerges from the billions of neurons that constitute the human brain is a major open problem in science that could bridge natural science –biology– to humanities –psychology. Psychology studies performed on humans with functional Magnetic Resonance Imaging (fMRI) can be used to probe the full repertoire of high-level cognitive functions. While analyzing the resulting image data for a given experiment is a relatively well-mastered process, the challenges in comparing data across multiple datasets poses serious limitation to the field. Indeed, such comparisons require to pool together brain images acquired under different settings and assess the effect of different *experimental conditions* that correspond to psychological effects studied by neuroscientists.

Such meta-analyses are now becoming possible thanks to the development of public data resources –OpenfMRI http://openfmri.org and NeuroVault http://neurovault.org. As many others, researchers of the Parietal team understand these data sources well and contribute to them. However, in such open-ended context, the description of experiments in terms of cognitive concepts is very difficult: there is no universal definition of cognitive terms that could be employed consistently by neuroscientists. Hence meta-analytic studies loose power and specificity. On the other hand, http://brainspell.org provide a set of curated annotation, albeit on much less data, that can serve as a seed or a ground truth to define a consensual ontology of cognitive concepts. Relating these terms to brain activity poses another challenge, of statistical nature, as brain patterns form high-dimensional data in perspective with the scarcity and the noise of the data.

The purpose of this project is to learn a semantic structure in cognitive terms from their occurrence in brain activation. This structure will simplify massive multi-label statistical-learning problems that arise in brain mapping by providing compact representations of cognitive concepts while capturing the imprecision on the definition these concepts.

9.1.4. HidimStat

Participants: Bertrand Thirion [Correspondant], Jerome-Alexis Chevalier, Joseph Salmon.

This project is funded by Digiteo.

This is a Digicosme project (2017-2020) and a collaboration with Joseph Salmon (Telecom Paritech).

The HiDimStat project aims at handling uncertainty in the challenging context of high dimensional regression problem. Though sparse models have been popularized in the last twenty years in contexts where many features can explain a phenomenon, it remains a burning issue to attribute confidence to the predictive models that they produce. Such a question is hard both from the statistical modeling point of view, and from a computation perspective. Indeed, in practical settings, the amount of features at stake (possibly up to several millions in high resolution brain imaging) limit the application of current methods and require new algorithms to achieve computational efficiency. We plan to leverage recent developments in sparse convex solvers as well as more efficient reformulations of testing and confidence interval estimates to provide several communities with practical software handling uncertainty quantification. Specific validation experiments will be performed in the field of brain imaging.

9.1.5. Template estimation for arbitrary alignments: application to brain imaging.

Participants: Bertrand Thirion [Correspondant], Thomas Bazeille.

This project is funded by Digiteo.

In the recent years, the nature of scientific inference has shifted quite substantially from model-based to predictive approaches, thanks to the generalization of powerful machine learning techniques. While this has certainly improved scientific standards, this has also obscured the objects and concepts on which inference is drawn. For instance, it is now possible –based on some initial data– to predict individual brain activity topographies, yet the very notion of a standard brain template has become increasingly elusive. Given the importance of establishing models for the progress of knowledge, we revisit the problem of model inference on data with high variance. Specifically, in a context where almost arbitrary transformation can successfully warp observations to each other with high accuracy, what is the common definition of a population model underlying all these observations? What is the working definition of a template ? We plan to leverage recent developments on optimal transport and multivariate analysis to build working definition of templates; we will use them in a brain imaging context to build a novel generation of brain templates.

9.1.6. AMPHI project

Participants: Bertrand Thirion [Correspondant], Joseph Salmon, Antonio Andre Monteiro Manoel.

This is a Digicosme project (2017-2020) and a collaboration with Joseph Salmon (Telecom Paritech) and Lenka Zdeborova (CEA, IPhT).

In many scientific fields, the data acquisition devices have benefited of hardware improvement to increase the resolution of the observed phenomena, leading to ever larger datasets. While the dimensionality has increased, the number of samples available is often limited, due to physical or financial limits. This is a problem when these data are processed with estimators that have a large sample complexity, such as multivariate statistical models. In that case it is very useful to rely on structured priors, so that the results reflect the state of knowledge on the phenomena of interest. The study of the human brain activity through neuroimaging belongs among these problems, with up to 10^6 features, yet a set of observations limited by cost and participant comfort. We are missing fast estimators for multivariate models with structured priors, that furthermore provide statistical control on the solution. Approximate message passing (AMP) methods are designed to work optimally with low- sample-complexity, they accommodate rather generic class of priors and come with an estimation of statistical significance. They are therefore well suited for our purposes. We want to join forces to design a new generation of inverse problem solvers that can take into account the complex structure of brain images and provide guarantees in the low-sample-complexity regime. To this end, we will first adapt AMP to the brain mapping setting, using first standard sparsity priors (e.g. Gauss-Bernoulli) on the model. We will then consider more complex structured priors that control the variation of the learned image patterns in space. Crucial gains are expected from the use of the EM algorithm for parameter setting, that comes naturally with AMP. We will also examine the estimators provided by AMP for statistical significance. AMPHI will design a reference inference toolbox released as a generic open source library. We expect a 3- to 10-fold improvement in CPU time, that will benefit to large-scale brain mapping investigations.

9.1.7. CDS2

Participants: Bertrand Thirion [Correspondant], Gaël Varoquaux, Guillaume Lemaitre, Joris Van Den Bossche.

CDS2 is an "Strategic research initiative" of the Paris Saclay University Idex http://datascience-paris-saclay.fr. Although it groups together many partners of the Paris Saclay ecosystem, Parietal has been deeply involved in the project. It currently funds a post-doc for Guillaume Lemaitre and an engineer positions for Joris van den Bossche. Alexandre Boucaud was funded till December as engineer.

9.2. National Initiatives

9.2.1. ANR

9.2.1.1. Neuroref: Mathematical Models of Anatomy / Neuroanatomy / Diffusion MRI

Participants: Demian Wassermann [Correspondant], Antonia Machlouzarides Shalit, Valentin Iovene.

While mild traumatic brain injury (mTBI) has become the focus of many neuroimaging studies, the understanding of mTBI, particularly in patients who evince no radiological evidence of injury and yet experience clinical and cognitive symptoms, has remained a complex challenge. Sophisticated imaging tools are needed to delineate the kind of subtle brain injury that is extant in these patients, as existing tools are often ill-suited for the diagnosis of mTBI. For example, conventional magnetic resonance imaging (MRI) studies have focused on seeking a spatially consistent pattern of abnormal signal using statistical analyses that compare average differences between groups, i.e., separating mTBI from healthy controls. While these methods are successful in many diseases, they are not as useful in mTBI, where brain injuries are spatially heterogeneous.

The goal of this proposal is to develop a robust framework to perform subject-specific neuroimaging analyses of Diffusion MRI (dMRI), as this modality has shown excellent sensitivity to brain injuries and can locate subtle brain abnormalities that are not detected using routine clinical neuroradiological readings. New algorithms will be developed to create Individualized Brain Abnormality (IBA) maps that will have a number of clinical and research applications. In this proposal, this technology will be used to analyze a previously acquired dataset from the INTRuST Clinical Consortium, a multi-center effort to study subjects with Post-Traumatic Stress Disorder (PTSD) and mTBI. Neuroimaging abnormality measures will be linked to clinical and neuropsychological assessments. This technique will allow us to tease apart neuroimaging differences between PTSD and mTBI and to establish baseline relationships between neuroimaging markers, and clinical and cognitive measures.

9.2.1.2. DirtyData: Data integration and cleaning for statistical analysis

Participants: Gaël Varoquaux [Correspondant], Patricio Cerda Reyes, Pierre Glaser.

Machine learning has inspired new markets and applications by extracting new insights from complex and noisy data. However, to perform such analyses, the most costly step is often to prepare the data. It entails correcting errors and inconsistencies as well as transforming the data into a single matrix-shaped table that comprises all interesting descriptors for all observations to study. Indeed, the data often results from merging multiple sources of informations with different conventions. Different data tables may come without names on the columns, with missing data, or with input errors such as typos. As a result, the data cannot be automatically shaped into a matrix for statistical analysis.

This proposal aims to drastically reduce the cost of data preparation by integrating it directly into the statistical analysis. Our key insight is that machine learning itself deals well with noise and errors. Hence, we aim to develop the methodology to do statistical analysis directly on the original dirty data. For this, the operations currently done to clean data before the analysis must be adapted to a statistical framework that captures errors and inconsistencies. Our research agenda is inspired from the data-integration state of the art in database research combined with statistical modeling and regularization from machine learning.

Data integrating and cleaning is traditionally performed in databases by finding fuzzy matches or overlaps and applying transformation rules and joins. To incorporate it in the statistical analysis, an thus propagate uncertainties, we want to revisit those logical and set operations with statistical-learning tools. A challenge is to turn the entities present in the data into representations well-suited for statistical learning that are robust to potential errors but do not wash out uncertainty.

Prior art developed in databases is mostly based on first-order logic and sets. Our project strives to capture errors in the input of the entries. Hence we formulate operations in terms of similarities. We address typing entries, deduplication -finding different forms of the same entity- building joins across dirty tables, and correcting errors and missing data.

Our goal is that these steps should be generic enough to digest directly dirty data without user-defined rules. Indeed, they never try to build a fully clean view of the data, which is something very hard, but rather include in the statistical analysis errors and ambiguities in the data.

The methods developed will be empirically evaluated on a variety of dataset, including the French public-data repository, data.gouv.fr. The consortium comprises a company specialized in data integration, Data Publica, that guides business strategies by cross-analyzing public data with market-specific data.

9.2.1.3. FastBig Project

Participants: Bertrand Thirion [Correspondant], Jerome-Alexis Chevalier, Tuan Binh Nguyen.

In many scientific applications, increasingly-large datasets are being acquired to describe more accurately biological or physical phenomena. While the dimensionality of the resulting measures has increased, the number of samples available is often limited, due to physical or financial limits. This results in impressive amounts of complex data observed in small batches of samples.

A question that arises is then : what features in the data are really informative about some outcome of interest ? This amounts to inferring the relationships between these variables and the outcome, conditionally to all other variables. Providing statistical guarantees on these associations is needed in many fields of data science, where competing models require rigorous statistical assessment. Yet reaching such guarantees is very hard.

FAST-BIG aims at developing theoretical results and practical estimation procedures that render statistical inference feasible in such hard cases. We will develop the corresponding software and assess novel inference schemes on two applications : genomics and brain imaging.

9.2.1.4. MultiFracs project

Participant: Philippe Ciuciu [Correspondant].

The scale-free concept formalizes the intuition that, in many systems, the analysis of temporal dynamics cannot be grounded on specific and characteristic time scales. The scale-free paradigm has permitted the relevant analysis of numerous applications, very different in nature, ranging from natural phenomena (hydrodynamic turbulence, geophysics, body rhythms, brain activity,...) to human activities (Internet traffic, population, finance, art,...).

Yet, most successes of scale-free analysis were obtained in contexts where data are univariate, homogeneous along time (a single stationary time series), and well-characterized by simple-shape local singularities. For such situations, scale-free dynamics translate into global or local power laws, which significantly eases practical analyses. Numerous recent real-world applications (macroscopic spontaneous brain dynamics, the central application in this project, being one paradigm example), however, naturally entail large multivariate data (many signals), whose properties vary along time (non-stationarity) and across components (non-homogeneity), with potentially complex temporal dynamics, thus intricate local singular behaviors.

These three issues call into question the intuitive and founding identification of scale-free to power laws, and thus make uneasy multivariate scale-free and multifractal analyses, precluding the use of univariate methodologies. This explains why the concept of scale-free dynamics is barely used and with limited successes in such settings and highlights the overriding need for a systematic methodological study of multivariate scale-free and multifractal dynamics. The Core Theme of MULTIFRACS consists in laying the theoretical foundations of a practical robust statistical signal processing framework for multivariate non homogeneous scale-free and multifractal analyses, suited to varied types of rich singularities, as well as in performing accurate analyses of scale-free dynamics in spontaneous and task-related macroscopic brain activity, to assess their natures, functional roles and relevance, and their relations to behavioral performance in a timing estimation task using multimodal functional imaging techniques.

This overarching objective is organized into 4 Challenges:

1. Multivariate scale-free and multifractal analysis,
2. Second generation of local singularity indices,
3. Scale-free dynamics, non-stationarity and non-homogeneity,
4. Multivariate scale-free temporal dynamics analysis in macroscopic brain activity.

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

9.3.1.1. Neurolang

Title: Accelerating Neuroscience Research by Unifying Knowledge Representation and Analysis Through a Domain Specific Language

Programm: ERC Starting researcher

Duration: 01/03/2018 - 28/02/2023

Coordinator: Demian Wassermann

Inria contact: Demian Wassermann

Summary:

Neuroscience is at an inflection point. The 150-year old cortical specialization paradigm, in which cortical brain areas have a distinct set of functions, is experiencing an unprecedented momentum with over 1000 articles being published every year. However, this paradigm is reaching its limits. Recent studies show that current approaches to atlas brain areas, like relative location, cellular population type, or connectivity, are not enough on their own to characterize a cortical area and its function unequivocally. This hinders the reproducibility and advancement of neuroscience.

Neuroscience is thus in dire need of a universal standard to specify neuroanatomy and function: a novel formal language allowing neuroscientists to simultaneously specify tissue characteristics, relative location, known function and connectional topology for the unequivocal identification of a given brain region.

The vision of NeuroLang is that a unified formal language for neuroanatomy will boost our understanding of the brain. By defining brain regions, networks, and cognitive tasks through a set of formal criteria, researchers will be able to synthesize and integrate data within and across diverse studies. NeuroLang will accelerate the development of neuroscience by providing a way to evaluate anatomical specificity, test current theories, and develop new hypotheses.

NeuroLang will lead to a new generation of computational tools for neuroscience research. In doing so, we will be shedding a novel light onto neurological research and possibly disease treatment and palliative care. Our project complements current developments in large multimodal studies across different databases. This project will bring the power of Domain Specific Languages to neuroscience research, driving the field towards a new paradigm articulating classical neuroanatomy with current statistical and machine learning-based approaches.

9.3.1.2. SLAB (698)

Title: Signal processing and Learning Applied to Brain data

Programm: ERC Starting researcher

Duration: 01/04/2017 - 31/08/2021

Coordinator: Alexandre Gramfort

Partner: LTCI , Telecom ParisTech (France)

Inria contact: Alexandre Gramfort

Summary:

Understanding how the brain works in healthy and pathological conditions is considered as one of the challenges for the 21st century. After the first electroencephalography (EEG) measurements in 1929, the 90's was the birth of modern functional brain imaging with the first functional MRI and full head magnetoencephalography (MEG) system. In the last twenty years, imaging has revolutionized clinical and cognitive neuroscience.

After pioneering works in physics and engineering, the field of neuroscience has to face two major challenges. The size of the datasets keeps growing. The answers to neuroscience questions are limited by the complexity of the signals observed: non-stationarity, high noise levels, heterogeneity of sensors, lack of accurate models. SLAB will provide the next generation of models and algorithms for mining electrophysiology signals which offer unique ways to image the brain at a millisecond time scale.

SLAB will develop dedicated machine learning and signal processing methods and favor the emergence of new challenges for these fields. SLAB focuses on five objectives: 1) source localization with M/EEG for brain imaging at high temporal resolution 2) representation learning to boost statistical power and reduce acquisition costs 3) fusion of heterogeneous sensors 4) modeling of non-stationary spectral interactions to identify functional coupling between neural ensembles 5) development of fast algorithms easy to use by non-experts.

SLAB aims to strengthen mathematical and computational foundations of brain data analysis. The methods developed will have applications across fields (computational biology, astronomy, econometrics). Yet, the primary impact of SLAB will be on neuroscience. The tools and high quality open software produced in SLAB will facilitate the analysis of electrophysiology data, offering new perspectives to understand how the brain works at a mesoscale, and for clinical applications (epilepsy, autism, tremor, sleep disorders).

9.3.1.3. Neuroimaging power (262)

Title: Neuroimaging power

Programm: Marie Curie Fellowhip

Duration: 01/11/2016 - 31/10/2019

Coordinator: Inria

Partner: BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY (United States)

Inria contact: Bertrand Thirion

Summary:

There is an increasing concern about statistical power in neuroscience research. Critically, an underpowered study has poor predictive power. Findings from a low-power study are unlikely to be reproducible, and thus a power analysis is a critical component of any paper. This project aims to promote and facilitate the use of power analyses.

A key component of a power analysis is the specification of an effect size. However, in neuroimaging, there is no standardised way to communicate effect sizes, which makes the choice of an appropriate effect size a formidable task. The best way today to perform a power analysis is by collecting a pilot data set, a very expensive practice. To eliminate the need for pilot data, we will develop a standardised measure of effect size taking into account the spatial variance and the uncertainty of the measurements. Communicating effect sizes in new publications will facilitate the use of power analyses.

To further alleviate the need for pilot data, we will provide a library of effect sizes for different tasks and contrasts, using open data projects in neuroimaging. We will integrate our effect size estimator in open repositories NeuroVault and OpenfMRI. Consequently, these effect sizes can then serve as a proxy for a pilot study, and as such, a huge cost in the design of an experiment is eliminated.

A new experiment will not be identical to the open data and as such the hypothesised parameters might not be fully accurate. To address this issue, we present a flexible framework to analyse data mid-way without harming the control of the type I error rate. Such a procedure will allow re-evaluating halfway an experiment whether it is useful to continue a study, and how many more subjects are needed for statistically sound inferences. To make our methods maximally available, we will write a software suite including all these methods in different programming platforms and we will provide a GUI to further increase the use of power analyses.

9.3.1.4. HBP SGA1

Title: Human Brain Project Specific Grant Agreement 1

Programm: FET Flagship

Duration: 01/04/2016 - 31/02/2020

Coordinator: Katrin Amunts

Partners: 150 european labs, please see https://www.humanbrainproject.eu/en/open-ethical-engaged/contributors/partners

Inria contact: Bertrand Thirion

Summary

Understanding the human brain is one of the greatest scientific challenges of our time. Such an understanding can provide profound insights into our humanity, leading to fundamentally new computing technologies, and transforming the diagnosis and treatment of brain disorders. Modern ICT brings this prospect within reach. The HBP Flagship Initiative (HBP) thus proposes a unique strategy that uses ICT to integrate neuroscience data from around the world, to develop a unified multi-level understanding of the brain and diseases, and ultimately to emulate its computational capabilities. The goal is to catalyze a global collaborative effort. During the HBP's first Specific Grant Agreement (SGA1), the HBP Core Project will outline the basis for building and operating a tightly integrated Research Infrastructure, providing HBP researchers and the scientific Community with unique resources and capabilities. Partnering Projects will enable independent research groups to expand the capabilities of the HBP Platforms, in order to use them to address otherwise intractable problems in neuroscience, computing and medicine in the future. In addition, collaborations with other national, European and international initiatives will create synergies, maximizing returns on research investment. SGA1 covers the detailed steps that will be taken to move the HBP closer to achieving its ambitious Flagship Objectives.

9.3.1.5. HBP SGA2

Title: Interactive Computing E-Infrastructure for the Human Brain Project

Programm: FET Flagship

Duration: 01/04/2018 - 31/03/2020

Coordinator: Katrin Amunts

Partners: see https://www.humanbrainproject.eu/en/open-ethical-engaged/contributors/partners/

Inria contact: Bertrand Thirion

Summary:

The HBP Flagship was launched by the European Commission's Future and Emerging Technologies (FET) scheme in October 2013, and is scheduled to run for ten years. The Flagships, represent a new partnering model for visionary, long-term European cooperative research in the European Research Area, demonstrating the potential for common research efforts. The HBP has the following main objectives:

- Create and operate a European scientific Research Infrastructure for brain research, cognitive neuroscience, and other brain-inspired sciences
- Gather, organise and disseminate data describing the brain and its diseases
- Simulate the brain
- Build multi-scale scaffold theory and models for the brain
- Develop brain-inspired computing, data analytics and robotics
- Ensure that the HBP's work is undertaken responsibly and that it benefits society.

More information on the HBP's Flagship Objectives is available in the Framework Partnership Agreement.

The timeline of the Project is split into multiple phases, each of which will be covered by a separate funding agreement. The current phase is Specific Grant Agreement Two (SGA2), which spans the two-year period from April 2018–April 2020. The HBP is funded via several sources. Total funding is planned to be in the region of EUR 1 billion; around one half of which will be provided by the European Union, and the other by Member States and private funding sources. The European Union contributed EUR 54 million to the Project in the Ramp-Up Phase (October 2013 to March 2016), EUR 89 million for the second phase (SGA1), and EUR 88 million for the current phase (SGA2). The FET Flagships Staff Working Document provides further information on how Flagships are funded.

9.4. International Initiatives

9.4.1. Inria International Labs

Inria@SiliconValley

Associate Team involved in the International Lab:

9.4.1.1. LargeBrainNets

Title: Characterizing Large-scale Brain Networks Using Novel Computational Methods for dMRI and fMRI-based Connectivity

International Partner (Institution - Laboratory - Researcher):

Stanford Cognitive & Systems Neuroscience Lab, Stanford Medical School, USA. Contact: Vinod Menon.

Start year: 2016

See also: http://www-sop.inria.fr/members/Demian.Wassermann/large-brain-nets.html

In the past two decades, brain imaging of neurotypical individuals and clinical populations has primarily focused on localization of function and structures in the brain, revealing activation in specific brain regions during performance of cognitive tasks through modalities such as functional MRI. In parallel, technologies to identify white matter structures have been developed using diffusion MRI. More recently, interest has shifted towards developing a deeper understanding of the brain's intrinsic architecture and its influence on cognitive and affective information processing. Using for this resting state fMRI and diffusion MRI to build the functional and structural networks of the human brain.

The human brain is a complex patchwork of interconnected regions, and graph-theoretical approaches have become increasingly useful for understanding how functionally connected systems engender, and constrain, cognitive functions. The functional nodes of the human brain and their structural inter-connectivity, collectively the "connectome", are, however, poorly understood. Critically, there is a dearth of computational methods for reliably identifying functional nodes of the brain and their structural inter-connectivity in vivo, despite an abundance of high-quality data from the Human Connectome Project (HCP). Devising and validating methods for investigating the human connectome has therefore taken added significance.

The first major goal of this project is to develop and validate appropriate sophisticated computational and mathematical tools for identifying functional nodes at the whole-brain level and measuring structural and functional connectivity between them, using state-of-the-art human brain imaging techniques and open-source HCP data. To this end, we will first develop and validate novel computational tools for (1) identifying stable functional nodes of the human brain using resting-state functional MRI and (2) measuring structural connectivity between functional nodes of the brain using multi-shell high-angular diffusion MRI. Due to the complementarity of the two imaging techniques fMRI and dMRI, our novel computational methods methods, the synergy between the two laboratories of this associate team will allow us to reveal in unprecedented detail the structural and functional connectivity of the human brain.

The second major goal of this project is to use our newly developed computational tools to characterize normal structural and functional brain networks in neurotypical adults.

Inria@SiliconValley

Associate Team involved in the International Lab:

9.4.1.2. Meta&Co

Title: Meta-Analysis of Neuro-Cognitive Associations

International Partner (Institution - Laboratory - Researcher):

Stanford (United States) - Psychology department. - Russel Poldrack

Start year: 2018

See also: http://team.inria.fr/parietal

Cognitive science and psychiatry describe mental operations: cognition, emotion, perception and their dysfunction. Cognitive neuroimaging bridge these mental concepts to their implementation in the brain, neural firing and wiring, by relying on functional brain imaging. Yet aggregating results from experiments probing brain activity into a consistent description faces the roadblock that cognitive concepts and brain pathologies are ill-defined. Separation between them is often blurry. In addition, these concepts and subdivisions may not correspond to actual brain structures or systems. To tackle this challenge, we propose to adapt data-mining techniques used to learn relationships in computational linguistics. Natural language processing uses distributional semantics to build semantic relationships and ontologies. New models are needed to learn relationships from heterogeneous signals: functional magnetic resonance images (fMRI), on the one hand, combined with related psychology and neuroimaging annotations or publications, on the other hand. Such a joint effort will rely on large publicly-available fMRI databases shared by Podrack Lab, as well as literature mining.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

- June 2018: Prof. Lilianne Mujica-Parodi (Univ Stony-Brook, NY USA)
- April-June 2018: Dr Abderrahim Halimi (Edinburgh, UK)
- October 2018: Prof. Nikos Makris (Harvard Medical School)

- December 2018: Dr. Lang Chen (Stanford Medical University)

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

Bertrand Thirion co-organized the "Imagerie médicale et apprentissage automatique : vers une intelligence artificielle ?" colloque at Collège de France on May, 2nd, 2018.

10.1.1.2. Member of the Organizing Committees

Demian Wassermann: MICCAI 2018

Gaël Varoquaux: PyParis 2018

10.1.2. Scientific Events Selection

10.1.2.1. Member of the Conference Program Committees

- Philippe Ciuciu: Chairman of scientific oral sessions in ISBI 2018 (Washington DC) and EUSIPCO 2018 (Roma)
- Bertrand Thirion: Member of Program Committee for the OHBM 2018 meeting (Singapore),
- Gaël Varoquaux: program committee of NIPS, ICML, ICLR.
- Alexandre Gramfort: program committee of NIPS, ICML, ICLR.
- Demian Wassermann: ISMRM 2018

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Philippe Ciuciu: Associate Editor for EUSPICO 2018, Roma
- Philippe Ciuciu: Associate Editor for ISBI 2019, Venice
- Alexandre Gramfort: Editor, NeuroImage, Journal of Machine Learning Research (JMLR), Frontiers in Brain Imaging Methods
- Bertrand Thirion: Editor, Frontiers in Brain Imaging Methods

10.1.3.2. Reviewer - Reviewing Activities

- Philippe Ciuciu is reviewer for Nature Communication, Biological Pysichiatry, Plos Computational Biology, Scientific Reports, Journal of Neuroscience, IEEE Trans Signal Processing, IEEE Signal Processing Letters, IEEE Trans Medical Imaging, Frontiers in Neuroscience, Magnetic Resonance in Medicine, SIAM Imaging Science
- Gaël Varoquaux: Nature Methods, JMLR, PLOS Bio, NeuroImage, IEEE TBME, IEEE TMI, Annals of Applied Statistics, Biological Psychiatry, MedIA, Science, GigaScience
- Alexandre Gramfort: JMLR, PLOS Computational Biology, NeuroImage, IEEE TBME, IEEE TMI, IEEE TSP, MedIA, NIPS, ICML, ICLR, ICASSP, Scientific Reports, Frontiers in Brain Imaging Methods, Journal of Neuroscience Methods
- Bertrand Thirion: Nature communications, Neuroimage, Medical Image Analysis, IEEE TMI, PNAS, PLOS Comp Bio, Brain Structure and Function, NIPS, ICML, IPMI, OHBM, PRNI, AISTATS

- Denis Engemann PLOS Biology, PLOS Computational Biology, Scientific Reports, Neuroimage, Neuroimage Clinical, Human Brain Mapping, Journal of Machine Learning Research, Brain Topography, Brain Connectivity, Journal of Alzheimer's Disease, Neuroscience of Consciousness, PLOS ONE, Frontiers in Neuroscience, Journal of Computational Neuroscience, Psychiatry and Clinical Neurosciences, Sensors
- Demian Wassermann: NeuroImage, MRM, JMRI, Brain Structure and Function, Cortex, MedIA.

10.1.4. Invited Talks

- Philippe Ciuciu, May 2018: Keynote Lecture at the 8th International Conference on New Computational Methods for Inverse Problems, "Distribution-controlled and optimally spread sampling trajectories for accelerated Magnetic Resonance Imaging" (http://complement.farman.ens-cachan.fr/documents_web_NCMIP_2018/NCMIP_2018_program.pdf)
- Philippe Ciuciu, Nov 2018: Heriott-Watt University, "SPARKLING: variable-density k-space filling curves for accelerated T2*-weighted MRI"
- Philippe Ciuciu, Dec 2018: SHFJ, CEA Orsay, "Recent Advances in Compressed Sensing MRI for Highly Accelerated T2* -weighted Imaging"
- Denis Engemann, Dec 2018, ICM, Paris, MEG user day, 'Preparing data for source reconstruction: dos and donts'
- Denis Engemann, Nov 2018, ICM, Paris, 'Machine Learning with MEG EEG in Cognitive Neurology. Challenges and Opportunities'
- Denis Engemann, Nov 2018, Paris Machine Learning Meetup, invited talk, 'Random forest methods for EEG-based diagnosis of disorders of consciousness'
- Denis Engemann, Oct 2018, CRNL, Lyon, invited talk, 'Large-Scale Analysis of MEG/EEG in Cognitive Neurology. Challenges and Opportunities'
- Denis Engemann, Oct 2018, CRNL, Lyon, 3-day MNE-Python training workshop
- Denis Engemann, Apr 2018, Hôpital Erasme, Brussels, invited talk, 'The Challenge of Large-Scale and Population Analysis using MEG/EEG' LCFC
- Alexandre Gramfort, Fév 2018, invited talk, séminaire du Centre de Mathématiques Appliquées (CMAP) de l'Ecole Polytechnique
- Alexandre Gramfort, Mars 2018, invited talk, conseil d'administration Institut National du Cancer (INCA), Paris
- Alexandre Gramfort, Mars 2018, invited talk, Center for Data Science, Grenoble
- Alexandre Gramfort, June 2018, invited talk, ICML workshop of reproducibility in machine learning, Stockholm
- Alexandre Gramfort, June 2018, oral presentations, OHBM Conference, Singapore
- Alexandre Gramfort, June 2018, invited talk, DTU, Copenhagen
- Alexandre Gramfort, Aug 2018, oral presentations, Biomag International Conference, Philadelphia
- Alexandre Gramfort, Sept 2018, invited talk, Imperial College, UK
- Alexandre Gramfort, Sept 2018, invited talk, BCG Gamma FreshFromTheLabs Conference, Paris
- Alexandre Gramfort, Oct 2018, invited talk, Universität Heidelberg, Heidelberg
- Alexandre Gramfort, Nov 2018, oral presentation, Society for Neuroscience conference, San Diego
- Alexandre Gramfort, Nov 2018, invited talk, France is AI conference, Paris
- Alexandre Gramfort, Dec 2018, invited talk, Montreal Artificial Intelligence and Neuroscience (MAIN) workshop, Montreal
- Bertrand Thirion, Jan 2018, IHES, Bures sur Yvette, 'Toward a rigorous statistical framework for brain mapping'

- Bertrand Thirion, June 2018, IPHT seminary, Gif sur Yvette, 'A rigorous causal framework for brain mapping'
- Bertrand Thirion, May 2018, High Tech Peripherique conference, 'Large-scale machine learning for medical imaging'
- Bertrand Thirion, April 2018, Tau seminary, Gif sur Yvette, 'Causal analysis for Brain Mapping'
- Bertrand Thirion, July 2018, JST workshop in Paris, 'Toward rigorous e-sciences: High-dimensional statistical inference'
- Bertrand Thirion, Dec 2018, ESSI colloquium, Evry, 'High-dimensional statistical inference for e-sciences'
- Demian Wassermann, Feb 2018, Stanford Medical School, 'Microstructure Imaging with Diffusion MRI'
- Demian Wassermann, June 2018, Harvard Medical School, 'Random Effect Models for Structure Connectivity-Based Cortical Clustering'
- Demian Wassermann, June 2018, New York University, 'Recent advances in Micro and Macro scale brain analysis with Diffusion MRI'
- Gaël Varoquaux, Sept 2018, IPAM (Institute for Pure and Applied Mathematics), UCLA, Long Program Science at Extreme Scales: Where Big Data Meets Large-Scale Computing
- Gaël Varoquaux, August 2018, keynote speaker NeuroInformatics, Montreal
- Gaël Varoquaux, August 2018, invited talk MILA (Montreal Institute for Learning Algorithms), Montreal
- Gaël Varoquaux, August 2018, invited talk MNI (Montréal Neurological Institute), Montréal
- Gaël Varoquaux, August 2018, invited talk Institut de Gériatrie, Montréal
- Gaël Varoquaux, June 2018, invited talk, BCG Gamma days, Paris
- Gaël Varoquaux, Nov 2018, invited talk, chair DAMI (Data, Analytics and Models for Insurance), Paris
- Gaël Varoquaux, Feb 2018, invited talk, Gatsby institute for theoretical neuroscience, UCL, London
- Gaël Varoquaux, Dev 2018, invited talk, Journée Nationale de la Science Ouverte, Paris
- Gaël Varoquaux, Sept 2018, invited talk, MASES International Workshop on Machine Learning and Software Engineering in Symbiosis, ASE (Automated Software Engineering), Montpellier
- Gaël Varoquaux, Sept 2018, invited talk, MICCAI workshop, Granada
- Gaël Varoquaux, June 2018, invited talk, Machine learning in the real world, Paris
- Gaël Varoquaux, July 2018, invited talk, DataIA–Japan Science and Technology agency, Paris
- Gaël Varoquaux, Feb 2018, invited talk, Pycon Belarus, Minsk
- Gaël Varoquaux, May 2018, invited talk, Brainhack Paris
- Gaël Varoquaux, Oct 2018, invited talk, Biomarker days, Toulouse

10.1.5. Leadership within the Scientific Community

- Demian Wassermann: Organising committee of the BrainHack community — Paris Chapter
- Demian Wassermann: ISMRM and MICCAI action organization towards reducing gender and minority biases.

10.1.6. Scientific Expertise

- Philippe Ciuciu has been member of the 45th ANR Scientic Evaluation Committee in charge of evaluating the projects dealing with maths, signal processing, computer science methods for medicine and biology.

- Alexandre Gramfort has been member of the 45th ANR Scientic Evaluation Committee (CE23) in charge of evaluating the projects dealing with machine learning, data knowledge, statistics, optimization.
- Demian Wassermann FET-OPEN ERCA Action
- Gaël Varoquaux was expert for the startup incubator agoranov

10.1.7. Research Administration

- In Nov 2018, Philippe Ciuciu Elected Vice-Chair of the SAT Biomedical Image & Signal Analytics (EURASIP technical committee) for 2019-2020.
- In Dec 2018, Philippe Ciuciu has been Elected IEEE Signal Processing Society Representative at the 2019 IEEE ISBI conference.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master: Alexandre Gramfort, Optimization for Data Science, 20h, Msc 2 Data Science Master Ecole Polytechnique, France

Master: Alexandre Gramfort, DataCamp, 20h, Msc 2 Data Science Master Ecole Polytechnique, France

Master: Alexandre Gramfort, Source Imaging with EEG and MEG, 7h, Msc 2 in Biomedical Imaging at Télécom Paristech

Doctoral School: Alexandre Gramfort, Sparse Optimization, 4h, school at iTWIST workshop, Marseille

Master: Bertrand Thirion, Functional neuroimaging and BCI, 12h, Master MVA, ENS Paris-Saclay, France

Master: Philippe, fMRI course: From acquisition to data Analysis, 6h, Msc 2 in Biomedical Imaging, Université Paris-Sud

Bachelor: Demian Wassermann, CSE201 class, 15h, C++ programming, Ecole Polytechnique

Extension: Demian Wassermann, Data Science, 20h, Ecole Polytechnique

Master: Gaël Varoquaux, Machine learning in Python, 3h, ENSAE

Master: Gaël Varoquaux, Functional brain connectivity, 7h, Msc 2 in Biomedical Imaging at Télécom Paristech

10.2.2. Supervision

The following PhD defense have taken place

HdR : Gaël Varoquaux, Estimating brain functional connectivity and its variations from fMRI, Université Paris VI, 25/03/2018

PhD : Stanislas Chambon, *Learning from electrophysiology time series: From scoring to event detection*, Université Paris-Saclay, 14/12/2018, under the direction of Alexandre Gramfort

PhD : Tom Dupré La Tour, *Non-linear models for neurophysiological time series*, Université Paris-Saclay, 26/11/2018, under the direction of Alexandre Gramfort

PhD : Yousra Bekhti, *Contributions to sparse source localization for MEG/EEG brain imaging*, Université Paris-Saclay, 22/03/2018, under the direction of Alexandre Gramfort

PhD : Mainak Jas, *Advances in automating analysis of neural time series data*, Université Paris-Saclay, 12/04/2018, under the direction of Alexandre Gramfort

PhD : Arthur Mensch, , Université Paris-Saclay, 30/09/2018, under the direction of Bertrand Thirion, Gaël Varoquaux and Julien Mairal.

PhD : Carole Lazarus, , Université Paris-Saclay, 30/09/2018, under the direction of Philippe Ciuciu.

PhD in progress : Pierre Ablin, , 01/10/2016, coadvised by Alexandre Gramfort and Jean-François Cardoso,

PhD in progress : Mathurin Massias, , 01/10/2016, coadvised by Alexandre Gramfort and Joseph Salmon,

PhD in progress : Hicham Janati, , 01/10/2017, , coadvised by Alexandre Gramfort and Marco Cuturi

PhD in progress : Quentin Bertrand, , 01/10/2018, coadvised by Alexandre Gramfort and Joseph Salmon

PhD in progress : Hubert Banville, , 01/10/2018, coadvised by Alexandre Gramfort and Denis Engemann

PhD in progress : David Sabbagh, , 01/10/2018, coadvised by Alexandre Gramfort and Denis Engemann

PhD in progress : Jérôme-Alexis Chevalier, , 01/10/2017, coadvised by Bertrand Thirion and Joseph Salmon

PhD in progress : Hugo Richard, , 01/10/2018, coadvised by Bertrand Thirion and Olivier Grisel

PhD in progress : Thomas Bazeille, , 01/10/2018, advised by Bertrand Thirion

PhD in progress : Tuan Binh Nguyen, , 01/10/2018, coadvised by Bertrand thirion and Sylvain Arlot

PhD in progress : Valentin Iovene, , 01/10/2018, advised by Demian Wassermann

PhD in progress : Antonia Machlouzarides Shalit, , 01/10/2018, coadvised by Demian Wassermann and Bertrand Thirion

PhD in progress : Loubna El Gueddari, , 01/10/2016, coadvised by Philippe Ciuciu and Alexandre Vignaud

PhD in progress : Hamza Cherkaoui, , 01/10/2017, advised by Philippe Ciuciu

PhD in progress : Patricio Cerda Reyes, , 01/10/2016, coadvised by Gaël Varoquaux and Balazs Kegl

PhD in progress : Maeliss Jallais, , 01/10/2018, advised by Demian Wassermann

PhD in progress : Jerome Dockès, , 01/10/2016, coadvised by Fabian Shuchanek and Gaël Varoquaux

10.2.3. Juries

Alexandre Gramfort has been involved in the following PhD committees:

- Marine Le Morvan (Reviewer) / Mines ParisTech
- Jérémy Guillon (Examiner) / Inria Paris
- Rémi Leblond (Examiner) / Inria Paris
- Andreas Trier (Reviewer) / DTU, Copenhague

Alexandre Gramfort was involved in a hiring committee for Telecom ParisTech and a hiring committee for Université Paris Diderot.

Bertrand Thirion has been involved in the following committees:

- PhD Sebastian Tarando / ESSI Evry
- PhD Guillermo Gallardo / Inria Sophia-Antipolis
- Habilitation Florent Meyniel / Université Paris VI

Bertrand Thirion was involved in a hiring committee for Centrale-Supelec.

Gaël Varoquaux was involved in a hiring committee for Centrale-Supelec and a hiring committee for Inria Saclay.

Demian Wassermann was involved in the following committees:

- PhD Guillermo Gallardo / Inria Sophia-Antipolis
- Habilitation Daniel Margullies / ICM – UPMC

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

- Philippe Ciuciu has been member of the Inria Saclay scientific commission since 2016
- Alexandre Gramfort is member of the steering committee of the Paris-Saclay Center for Data Science
- Alexandre is a Member of the technical development committee of Inria Saclay since Dec 2018
- Bertrand Thirion is Deputy head of reasearch of Inria Saclay research center
- Bertrand Thirion, Leader of the Datasense research axis of the Digicosme Labex
- Bertrand Thirion, Member of the steering committee of the Dataia Convergence Institute
- Bertrand Thirion, Member of the steering committee of the Computer Science Department of Paris Saclay University.
- Gaël Varoquaux was Member of the technical development committee of Inria Saclay until Dec 2018
- Gaël Varoquaux is member of the doctoral monitoring committee of Inria Saclay
- Gaël Varoquaux is member of the steering committee of the Paris-Saclay Center for Data Science

10.3.2. Articles and contents

- Gaël Varoquaux was interviewed by ActuIA
- Alexandre Gramfort was interviewed by ActuIA
- Alexandre Gramfort was interviewed by Libération
- Alexandre Gramfort, Olivier Grisel Gaël Varoquaux are interviewed by Les Echos

10.3.3. Interventions

- Olivier Grisel gave an interview at the Paris Open Source Summit

10.3.4. Internal action

- In June 2018, Philippe Ciuciu did a training session on the DRF Impulsion funding mechanism at the seminar of new hired CEA DRF scientists"
- In April 2018, Denis Engemann and Alexandre Gramfort gave a 2-day educational course for cognitive neuroscientists at NeuroSpin on analysis of MEG and EEG data using Python.

10.3.5. Creation of media or tools for science outreach

Philippe Ciuciu made Two videos together with the CEA communication division on COSMIC project (funded by CEA)

- Long version: https://www.youtube.com/watch?v=gQh6D_vpkSo&t=23s&fbclid=IwAR0OcU3JEy4KQmo6DD-iN8otjlowyeXDkEF7ljniuzOkI-aWhGqyHhQzruo
- Short version: https://www.youtube.com/watch?v=p_KMEQGK-WA&fbclid=IwAR3xcf8e98M77lF9mpQSDWfKppjOt7p

Gaël Varoquaux, Olivier Grisel, Alexandre Gramfort, Guillaume Lemaître, Joris van den Bossche participated in a general-public movie about scikit-learn: https://www.youtube.com/watch?v=twqdXTCkeyk&t=9s

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] A. MENSCH.*Learning representations from functional MRI data*, Université Paris-Saclay, September 2018, https://tel.archives-ouvertes.fr/tel-01891633

Articles in International Peer-Reviewed Journal

[2] P. A. ABLIN, J.-F. CARDOSO, A. GRAMFORT.*Faster Independent Component Analysis by Preconditioning With Hessian Approximations*, in "IEEE Transactions on Signal Processing", August 2018, vol. 66, n^o 15, p. 4040-4049, https://hal.inria.fr/hal-01970746

[3] Y. BEKHTI, F. LUCKA, J. SALMON, A. GRAMFORT.*A hierarchical Bayesian perspective on majorization-minimization for non-convex sparse regression: application to M/EEG source imaging*, in "Inverse Problems", August 2018, vol. 34, n^o 8, 085010, https://hal.inria.fr/hal-01970744

[4] D. BZDOK, N. ALTMAN, M. KRZYWINSKI.*Points of Significance: Statistics versus Machine Learning*, in "Nature Methods", April 2018, p. 1-7, https://hal.archives-ouvertes.fr/hal-01723223

[5] D. BZDOK, M. KRZYWINSKI, N. ALTMAN.*Machine learning: Supervised methods, SVM and kNN*, in "Nature Methods", January 2018, p. 1-6, https://hal.archives-ouvertes.fr/hal-01657491

[6] D. BZDOK, A. MEYER-LINDENBERG.*Machine learning for precision psychiatry: Opportunites and challenges*, in "Biological Psychiatry: Cognitive Neuroscience and Neuroimaging", February 2018, https://hal.archives-ouvertes.fr/hal-01643933

[7] P. CERDA, G. VAROQUAUX, B. KÉGL.*Similarity encoding for learning with dirty categorical variables*, in "Machine Learning", June 2018, https://arxiv.org/abs/1806.00979 [*DOI :* 10.1007/S10994-018-5724-2], https://hal.inria.fr/hal-01806175

[8] S. CHAMBON, M. GALTIER, P. J. ARNAL, G. WAINRIB, A. GRAMFORT.*A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series*, in "IEEE Transactions on Neural Systems and Rehabilitation Engineering", March 2018, vol. 26, n^o 4, 17683810, https://arxiv.org/abs/1707.03321 [*DOI :* 10.1109/TNSRE.2018.2813138], https://hal.archives-ouvertes.fr/hal-01810436

[9] C. CURY, J. GLAUNÈS, R. TORO, M. CHUPIN, G. SCHUMANN, V. FROUIN, J.-B. POLINE, O. COLLIOT.*Statistical Shape Analysis of Large Datasets Based on Diffeomorphic Iterative Centroids*, in "Frontiers in Neuroscience", November 2018, vol. 12 [*DOI :* 10.3389/FNINS.2018.00803], https://hal.inria.fr/hal-01920263

[10] E. DOHMATOB, G. VAROQUAUX, B. THIRION.*Inter-subject registration of functional images: do we need anatomical images ?*, in "Frontiers in Neuroscience", March 2018, https://hal.archives-ouvertes.fr/hal-01701619

[11] D. ENGEMANN, F. RAIMONDO, J.-R. KING, B. ROHAUT, G. LOUPPE, F. FAUGERAS, J. ANNEN, H. CASSOL, O. GOSSERIES, D. FERNANDEZ-SLEZAK, S. LAUREYS, L. NACCACHE, S. DEHAENE, J.

SITT.*Robust EEG-based cross-site and cross-protocol classification of states of consciousness*, in "Brain - A Journal of Neurology ", October 2018, vol. 141, n^{o} 11, p. 3179–3192 [*DOI :* 10.1093/BRAIN/AWY251], https://hal.inria.fr/hal-01887793

[12] P. FILIPIAK, R. H. FICK, A. PETIET, M. SANTIN, A.-C. PHILIPPE, S. LEHÉRICY, P. CIUCIU, R. DERICHE, D. WASSERMANN.*Reducing the number of samples in spatiotemporal dMRI acquisition design*, in "Magnetic Resonance in Medicine", November 2018 [*DOI :* 10.1002/MRM.27601], https://hal.archives-ouvertes.fr/hal-01928734

[13] F. HADJ-SELEM, T. LOFSTEDT, E. DOHMATOB, V. FROUIN, M. DUBOIS, V. GUILLEMOT, E. DUCHESNAY.*Continuation of Nesterov's Smoothing for Regression with Structured Sparsity in High-Dimensional Neuroimaging*, in "IEEE Transactions on Medical Imaging", 2018, vol. 2018 [*DOI :* 10.1109/TMI.2018.2829802], https://hal-cea.archives-ouvertes.fr/cea-01883286

[14] G. HARTWIGSEN, D. BZDOK.*Multivariate single-subject analysis of short-term reorganization in the language network*, in "Cortex", July 2018, 4 [*DOI :* 10.1016/J.CORTEX.2018.06.013], https://hal.archives-ouvertes.fr/hal-01824229

[15] Y. HONG, L. J. O'DONNELL, P. SAVADJIEV, F. ZHANG, D. WASSERMANN, O. PASTERNAK, H. J. JOHNSON, J. PAULSEN, J.-P. VONSATTEL, N. MAKRIS, C.-F. WESTIN, Y. RATHI.*Genetic load determines atrophy in hand cortico-striatal pathways in presymptomatic Huntington's disease*, in "Human Brain Mapping", 2018 [*DOI :* 10.1002/HBM.24217], https://hal.inria.fr/hal-01787886

[16] M. JAS, E. . LARSON, D. ENGEMANN, J. LEPPÄKANGAS, S. TAULU, M. HÄMÄLÄINEN, A. GRAMFORT.*A Reproducible MEG/EEG Group Study With the MNE Software: Recommendations, Quality Assessments, and Good Practices*, in "Frontiers in Neuroscience", August 2018, vol. 12 [*DOI :* 10.3389/FNINS.2018.00530], https://hal.archives-ouvertes.fr/hal-01854552

[17] J. M. KERNBACH, T. D. SATTERTHWAITE, D. S. BASSETT, J. SMALLWOOD, D. MARGULIES, S. KRALL, P. SHAW, G. VAROQUAUX, B. THIRION, K. KONRAD, D. BZDOK.*Shared Endo-phenotypes of Default Mode Dysfunction in Attention Deficit/Hyperactivity Disorder and Autism Spectrum Disorder*, in "Translational Psychiatry", July 2018, https://hal.archives-ouvertes.fr/hal-01790245

[18] J. M. KERNBACH, B. T. T. YEO, J. SMALLWOOD, D. MARGULIES, M. THIEBAUT DE SCHOTTEN, H. WALTER, M. SABUNCU, A. J. HOLMES, A. GRAMFORT, G. VAROQUAUX, B. THIRION, D. BZDOK.*Subspecialization within default mode nodes characterized in 10,000 UK Biobank participants*, in "Proceedings of the National Academy of Sciences of the United States of America ", November 2018 [*DOI :* 10.1073/PNAS.1804876115], https://hal.archives-ouvertes.fr/hal-01926796

[19] M. KOWALSKI, A. MEYNARD, H.-T. WU.*Convex Optimization approach to signals with fast varying instantaneous frequency*, in "Applied and Computational Harmonic Analysis", January 2018, vol. 44, n^{o} 1, p. 89 - 122, https://arxiv.org/abs/1503.07591 [*DOI :* 10.1016/J.ACHA.2016.03.008], https://hal.archives-ouvertes.fr/hal-01199615

[20] C. LAZARUS, P. WEISS, A. VIGNAUD, P. CIUCIU.*An Empirical Study of the Maximum Degree of Undersampling in Compressed Sensing for T2*-weighted MRI*, in "Magnetic Resonance Imaging", 2018, p. 1-31, https://hal.inria.fr/hal-01829323

[21] J. LEFORT-BESNARD, D. S. BASSETT, J. SMALLWOOD, D. S. MARGULIES, B. DERNTL, O. GRUBER, A. ALEMAN, R. JARDRI, G. VAROQUAUX, B. THIRION, S. B. EICKHOFF, D. BZDOK.*Different shades of default mode disturbance in schizophrenia: Subnodal covariance estimation in structure and function*, in "Human Brain Mapping", January 2018, p. 1-52, https://hal.archives-ouvertes.fr/hal-01620441

[22] J. LEFORT-BESNARD, G. VAROQUAUX, B. DERNTL, O. GRUBER, A. ALEMAN, R. JARDRI, I. SOMMER, B. THIRION, D. BZDOK.*Patterns of Schizophrenia Symptoms: Hidden Structure in the PANSS Questionnaire*, in "Translational Psychiatry", 2018, https://hal.archives-ouvertes.fr/hal-01888918

[23] L. M. M. LÊ, B. KÉGL, A. GRAMFORT, C. MARINI, D. NGUYEN, M. CHERTI, S. TFAILI, A. TFAYLI, A. BAILLET-GUFFROY, P. PROGNON, P. CHAMINADE, E. CAUDRON.*Optimization of classification and regression analysis of four monoclonal antibodies from Raman spectra using collaborative machine learning approach*, in "Talanta", July 2018, vol. 184, p. 260-265, https://hal.archives-ouvertes.fr/hal-01969999

[24] A. MENSCH, J. MAIRAL, B. THIRION, G. VAROQUAUX.*Stochastic Subsampling for Factorizing Huge Matrices*, in "IEEE Transactions on Signal Processing", January 2018, vol. 66, n^{o} 1, p. 113-128, https://arxiv.org/abs/1701.05363 [*DOI :* 10.1109/TSP.2017.2752697], https://hal.archives-ouvertes.fr/hal-01431618

[25] A. L. PINHO, A. AMADON, T. RUEST, M. FABRE, E. DOHMATOB, I. DENGHIEN, C. GINISTY, S. BECUWE-DESMIDT, S. ROGER, L. LAURIER, V. JOLY-TESTAULT, G. MÉDIOUNI-CLOAREC, C. DOUBLÉ, B. MARTINS, P. PINEL, E. EGER, G. VAROQUAUX, C. PALLIER, S. DEHAENE, L. HERTZ-PANNIER, B. THIRION.*Individual Brain Charting, a high-resolution fMRI dataset for cognitive mapping*, in "Scientific Data ", June 2018, vol. 5, 180105 [*DOI :* 10.1038/SDATA.2018.105], https://hal.archives-ouvertes.fr/hal-01817528

[26] V. SYDNOR, A. M. RIVAS-GRAJALES, A. LYALL, F. ZHANG, S. BOUIX, S. KARMACHARYA, M. SHENTON, C.-F. WESTIN, N. MAKRIS, D. WASSERMANN, L. J. O'DONNELL, M. KUBICKI.*A comparison of three fiber tract delineation methods and their impact on white matter analysis*, in "NeuroImage", May 2018, vol. 178, p. 318-331 [*DOI :* 10.1016/J.NEUROIMAGE.2018.05.044], https://hal.inria.fr/hal-01807178

[27] G. VAROQUAUX, R. POLDRACK.*Predictive models avoid excessive reductionism in cognitive neuroimaging*, in "Current Opinion in Neurobiology", April 2019, vol. 55 [*DOI :* 10.1016/J.CONB.2018.11.002], https://hal.archives-ouvertes.fr/hal-01856412

[28] G. VAROQUAUX, Y. SCHWARTZ, R. POLDRACK, B. GAUTHIER, D. BZDOK, J.-B. POLINE, B. THIRION.*Atlases of cognition with large-scale brain mapping*, in "PLoS Computational Biology", 2018, https://hal.inria.fr/hal-01908189

[29] L. WALLER, A. BROVKIN, L. DORFSCHMIDT, D. BZDOK, H. WALTER, J. D. KRUSCHWITZ.*GraphVar 2.0: a user-friendly toolbox for machine learning on functional connectivity measures*, in "Journal of Neuroscience Methods", January 2018, 40, https://hal.archives-ouvertes.fr/hal-01828991

[30] H.-T. WANG, D. BZDOK, D. MARGULIES, C. CRADDOCK, M. MILHAM, E. JEFFERIES, J. SMALLWOOD.*Patterns of thought: population variation in the associations between large-scale network organisation and self-reported experiences at rest*, in "NeuroImage", May 2018, https://hal.archives-ouvertes.fr/hal-01782292

[31] A. DE PIERREFEU, T. FOVET, F. HADJ-SELEM, T. LOFSTEDT, P. CIUCIU, S. LEFEBVRE, P. THOMAS, R. LOPES, R. JARDRI, E. DUCHESNAY.*Prediction of activation patterns preceding hallucinations in patients

with schizophrenia using machine learning with structured sparsity, in "Human Brain Mapping", April 2018, vol. 39, n^o 4, p. 1777 - 1788 [*DOI :* 10.1002/HBM.23953], https://hal-cea.archives-ouvertes.fr/cea-01883271

[32] A. DE PIERREFEU, T. LOFSTEDT, F. HADJ-SELEM, M. DUBOIS, R. JARDRI, T. FOVET, P. CIUCIU, V. FROUIN, E. DUCHESNAY.*Structured Sparse Principal Components Analysis With the TV-Elastic Net Penalty*, in "IEEE Transactions on Medical Imaging", February 2018, vol. 37, n^o 2, p. 396 - 407 [*DOI :* 10.1109/TMI.2017.2749140], https://hal-cea.archives-ouvertes.fr/cea-01883278

[33] A. DE PIERREFEU, T. LÖFSTEDT, C. LAIDI, F. HADJ-SELEM, J. BOURGIN, T. HAJEK, F. SPANIEL, M. KOLENIC, P. CIUCIU, N. HAMDANI, M. LEBOYER, T. FOVET, R. JARDRI, J. HOUENOU, E. DUCHESNAY.*Identifying a neuroanatomical signature of schizophrenia, reproducible across sites and stages, using machine learning with structured sparsity*, in "Acta Psychiatrica Scandinavica", 2018, vol. 2018, p. 1 - 10 [*DOI :* 10.1111/ACPS.12964], https://hal-cea.archives-ouvertes.fr/cea-01883283

International Conferences with Proceedings

[34] P. A. ABLIN, J.-F. CARDOSO, A. GRAMFORT.*Accelerating likelihood optimization for ICA on real signals*, in "LVA-ICA 2018", Guildford, United Kingdom, July 2018, https://arxiv.org/abs/1806.09390 , https://hal.inria.fr/hal-01822602

[35] A. ALIMI, R. H. FICK, D. WASSERMANN, R. DERICHE.*Dmipy, a Diffusion Microstructure Imaging toolbox in Python to improve research reproducibility*, in "MICCAI 2018 - Workshop on Computational Diffusion MRI", Granada, Spain, September 2018, https://hal.inria.fr/hal-01873353

[36] H. CHERKAOUI, L. E. GUEDDARI, C. LAZARUS, A. GRIGIS, F. POUPON, A. VIGNAUD, S. FARRENS, J.-L. STARCK, P. CIUCIU.*Analysis vs Synthesis-based Regularization for combined Compressed Sensing and Parallel MRI Reconstruction at 7 Tesla*, in "26th European Signal Processing Conference (EUSIPCO 2018)", Roma, Italy, September 2018, https://hal.inria.fr/hal-01800700

[37] J. DOCKÈS, D. WASSERMANN, R. POLDRACK, F. M. SUCHANEK, B. THIRION, G. VAROQUAUX.*Text to brain: predicting the spatial distribution of neuroimaging observations from text reports*, in "MICCAI 2018 - 21st International Conference on Medical Image Computing and Computer Assisted Intervention", Granada, Spain, September 2018, p. 1-18, https://arxiv.org/abs/1806.01139 , https://hal.archives-ouvertes.fr/hal-01807295

[38] L. EL GUEDDARI, C. LAZARUS, H. CARRIÉ, A. VIGNAUD, P. CIUCIU.*Self-calibrating nonlinear reconstruction algorithms for variable density sampling and parallel reception MRI*, in "10th IEEE Sensor Array and Multichannel Signal Processing workshop", Sheffield, United Kingdom, July 2018, p. 1-5, https://hal.inria.fr/hal-01782428

[39] P. FILIPIAK, R. FICK, A. PETIET, M. SANTIN, A.-C. PHILIPPE, S. LEHÉRICY, R. DERICHE, D. WASSERMANN.*Coarse-Grained Spatiotemporal Acquisition Design for Diffusion MRI*, in "ISBI 2019 - Proceedings of The IEEE International Symposium on Biomedical Imaging", Venice, Italy, April 2019, https://hal.inria.fr/hal-01973588

[40] G. GALLARDO, N. GAYRAUD, R. DERICHE, M. CLERC, S. DESLAURIERS-GAUTHIER, D. WASSERMANN.*Solving the Cross-Subject Parcel Matching Problem using Optimal Transport*, in "International Conference on Medical Image Computing and Computer-Assisted Intervention 2018", Granada, Spain, September 2018, https://hal.archives-ouvertes.fr/hal-01935684

[41] D. LA ROCCA, P. CIUCIU, V. VAN WASSENHOVE, H. WENDT, P. ABRY, R. LEONARDUZZI.*Scale-free functional connectivity analysis from source reconstructed MEG data*, in "EUSIPCO 2018 - 26th European Signal Processing Conference", Roma, Italy, September 2018, p. 1-5, https://hal.inria.fr/hal-01800620

[42] C. MAUMET, G. FLANDIN, M. PEREZ-GUEVARA, J.-B. POLINE, J. RAJENDRA, R. REYNOLDS, B. THIRION, T. E. NICHOLS.*A standardised representation for non-parametric fMRI results*, in "OHBM 2018 - Annual meeting of the Organization of Human Brain Mapping", Singapore, Singapore, June 2018, p. 1-4, http://www.hal.inserm.fr/inserm-01828914

[43] A. MENSCH, M. BLONDEL.*Differentiable Dynamic Programming for Structured Prediction and Attention*, in "35th International Conference on Machine Learning", Stockholm, Sweden, Proceedings of the 35th International Conference on Machine Learning, July 2018, vol. 80, https://hal.archives-ouvertes.fr/hal-01809550

[44] H. WENDT, P. ABRY, P. CIUCIU.*Spatially regularized wavelet leader scale-free analysis of fMRI data*, in "IEEE International Symposium on Biomedical Imaging", Washington, DC, United States, April 2018, https://hal.inria.fr/hal-01782332

Conferences without Proceedings

[45] S. CHAMBON, V. THOREY, P. J. ARNAL, E. MIGNOT, A. GRAMFORT.*A deep learning architecture to detect events in EEG signals during sleep*, in "IEEE International Workshop on Machine Learning for Signal Processing (MLSP 2018)", Aalborg, Denmark, September 2018, https://hal.archives-ouvertes.fr/hal-01917529

[46] D. CHYZHYK, G. VAROQUAUX, B. THIRION, M. MILHAM.*Controlling a confound in predictive models with a test set minimizing its effect*, in "PRNI 2018 - 8th International Workshop on Pattern Recognition in Neuroimaging", Singapore, Singapore, June 2018, p. 1-4, https://hal.archives-ouvertes.fr/hal-01831701

[47] T. DUPRÉ LA TOUR, Y. GRENIER, A. GRAMFORT.*Driver estimation in non-linear autoregressive models*, in "43nd IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018)", Calgary, Canada, April 2018, https://hal.archives-ouvertes.fr/hal-01696786

[48] M. FRIGO, G. GALLARDO, I. COSTANTINI, A. DADUCCI, D. WASSERMANN, R. DERICHE, S. DESLAURIERS-GAUTHIER.*Reducing false positive connection in tractograms using joint structure-function filtering*, in "OHBM 2018 - Organization for Human Brain Mapping", Singapore, Singapore, June 2018, p. 1-3, https://hal.inria.fr/hal-01737434

[49] G. GALLARDO, S. BOUIX, D. WASSERMANN.*Diffusion Driven Label Fusion for White Matter Multi-Atlas Segmentation*, in "OHBM 2018 - Organization for Human Brain Mapping", Singapore, Singapore, June 2018, p. 1-2, https://hal.archives-ouvertes.fr/hal-01737422

[50] N. T. H. GAYRAUD, G. GALLARDO, M. CLERC, D. WASSERMANN.*Solving the Cross-Subject Parcel Matching Problem: Comparing Four Methods Using Extrinsic Connectivity*, in "OHBM 2018", Singapore, Singapore, June 2018, https://hal.archives-ouvertes.fr/hal-01737366

[51] T. KERDREUX, F. PEDREGOSA, A. D'ASPREMONT.*Frank-Wolfe with Subsampling Oracle*, in "ICML 2018 - 35th International Conference on Machine Learning", Stockholm, Sweden, July 2018, https://arxiv.org/abs/1803.07348 , https://hal.archives-ouvertes.fr/hal-01927391

[52] T. D. La Tour, T. Moreau, M. Jas, A. Gramfort.*Multivariate Convolutional Sparse Coding for Electromagnetic Brain Signals*, in "Advances in Neural Information Processing Systems (NeurIPS)", Montréal, Canada, December 2018, https://arxiv.org/abs/1805.09654 , https://hal.archives-ouvertes.fr/hal-01966685

[53] M. Massias, O. Fercoq, A. Gramfort, J. Salmon.*Generalized Concomitant Multi-Task Lasso for Sparse Multimodal Regression*, in "21st International Conference on Artificial Intelligence and Statistics (AISTATS 2018)", Lanzarote, Spain, April 2018, https://hal.archives-ouvertes.fr/hal-01812011

[54] M. Massias, A. Gramfort, J. Salmon.*Celer: a Fast Solver for the Lasso with Dual Extrapolation*, in "ICML 2018 - 35th International Conference on Machine Learning", Stockholm, Sweden, PMLR, July 2018, vol. 80, p. 3321-3330, https://hal.archives-ouvertes.fr/hal-01833398

[55] H. Richard, A. Pinho, B. Thirion, G. Charpiat.*Optimizing deep video representation to match brain activity*, in "CCN 2018 - Conference on Cognitive Computational Neuroscience", Philadelphia, United States, September 2018, https://arxiv.org/abs/1809.02440 , https://hal.archives-ouvertes.fr/hal-01868735

[56] J.-B. Schiratti, J.-E. Le Douget, M. Le Van Quyen, S. Essid, A. Gramfort.*An ensemble learning approach to detect epileptic seizures from long intracranial EEG recordings*, in "International Conference on Acoustics, Speech, and Signal Processing", Calgary, Canada, April 2018, https://hal.archives-ouvertes.fr/hal-01724272

[57] A. de Pierrefeu, T. Lofstedt, C. Laidi, F. Hadj-Selem, M. Leboyer, P. Ciuciu, J. Houenou, E. Duchesnay.*Interpretable and stable prediction of schizophrenia on a large multisite dataset using machine learning with structured sparsity*, in "2018 International Workshop on Pattern Recognition in Neuroimaging (PRNI)", Singapore, Singapore, IEEE, June 2018 [*DOI :* 10.1109/PRNI.2018.8423946], https://hal-cea.archives-ouvertes.fr/cea-01883311

Other Publications

[58] P. A. Ablin, J.-F. Cardoso, A. Gramfort.*Beyond Pham's algorithm for joint diagonalization*, November 2018, https://arxiv.org/abs/1811.11433 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01936887

[59] P. A. Ablin, D. Fagot, H. Wendt, A. Gramfort, C. Févotte.*A Quasi-Newton algorithm on the orthogonal manifold for NMF with transform learning*, November 2018, https://arxiv.org/abs/1811.02225 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01912918

[60] F. Almairac, P. Filipiak, L. Slabu, M. Clerc, T. Papadopoulo, D. Fontaine, L. Mondot, S. Chanelet, D. Wassermann, R. Deriche.*Bridging Brain Structure and Function by Correlating Structural Connectivity and Cortico-Cortical Transmission*, June 2018, 2nd C@UCA meeting, Poster, https://hal.inria.fr/hal-01852956

[61] D. Bzdok, D. Engemann, O. Grisel, G. Varoquaux, B. Thirion.*Prediction and inference diverge in biomedicine: Simulations and real-world data*, April 2018, working paper or preprint [*DOI :* 10.1101/327437], https://hal.archives-ouvertes.fr/hal-01848319

[62] D. Bzdok, T. M. Karrer.*Single-Subject Prediction: A Statistical Paradigm for Precision Psychiatry*, February 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01714822

[63] K. Dadi, M. Rahim, A. Abraham, D. Chyzhyk, M. Milham, B. Thirion, G. Varoquaux.*Benchmarking functional connectome-based predictive models for resting-state fMRI*, November 2018, working paper or preprint, https://hal.inria.fr/hal-01824205

[64] M. Gabrié, A. Manoel, C. Luneau, J. Barbier, N. Macris, F. Krzakala, L. Zdeborová.*Entropy and mutual information in models of deep neural networks*, November 2018, working paper or preprint, https://hal-cea.archives-ouvertes.fr/cea-01930228

[65] J.-R. King, L. . Gwilliams, C. . Holdgraf, J. . Sassenhagen, A. . Barachant, D. . Engemann, E. . Larson, A. Gramfort.*Encoding and Decoding Neuronal Dynamics: Methodological Framework to Uncover the Algorithms of Cognition*, July 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01848442

[66] C. Lazarus, P. Weiss, N. Chauffert, F. Mauconduit, L. El Gueddari, C. Destrieux, I. Zemmoura, A. Vignaud, P. Ciuciu.*Variable-density k-space filling curves for accelerated Magnetic Resonance Imaging*, August 2018, working paper or preprint, https://hal.inria.fr/hal-01861760

[67] A. Manoel, F. Krzakala, B. Thirion, G. Varoquaux, L. Zdeborová.*Approximate message-passing for convex optimization with non-separable penalties*, November 2018, https://arxiv.org/abs/1809.06304 - working paper or preprint, https://hal-cea.archives-ouvertes.fr/cea-01932983

[68] A. Mensch, J. Mairal, B. Thirion, G. Varoquaux.*Extracting Universal Representations of Cognition across Brain-Imaging Studies*, October 2018, https://arxiv.org/abs/1809.06035 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01874713

[69] D. Wassermann, D. V. Nguyen, G. Gallardo, J.-R. Li, W. Cai, V. Menon.*Sensing Von Economo Neurons in the Insula with Multi-shell Diffusion MRI*, 2018, International Society for Magnetic Resonance in Medicine, Poster, https://hal.inria.fr/hal-01807704

References in notes

[70] K. S. Button, J. P. Ioannidis, C. Mokrysz, B. A. Nosek, J. Flint, E. S. Robinson, M. R. Munafò.*Power failure: why small sample size undermines the reliability of neuroscience*, in "Nature Reviews Neuroscience", 2013, vol. 14, n^{o} 5, p. 365–376

[71] R. A. Poldrack, C. I. Baker, J. Durnez, K. J. Gorgolewski, P. M. Matthews, M. R. Munafò, T. E. Nichols, J.-B. Poline, E. Vul, T. Yarkoni.*Scanning the horizon: towards transparent and reproducible neuroimaging research*, in "Nature Reviews Neuroscience", 2017, vol. 18, n^{o} 2, p. 115–126

Project-Team PARSIFAL

Proof search and reasoning with logic specifications

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

IN PARTNERSHIP WITH:
CNRS

Ecole Polytechnique

RESEARCH CENTER
Saclay - Île-de-France

THEME
Proofs and Verification

Table of contents

Project-Team PARSIFAL

Creation of the Project-Team: 2007 July 01

Keywords:

Computer Science and Digital Science:

A2.1. - Programming Languages
A2.1.1. - Semantics of programming languages
A2.1.4. - Functional programming
A2.1.5. - Constraint programming
A2.1.10. - Domain-specific languages
A2.2.1. - Static analysis
A2.4.3. - Proofs
A2.5.4. - Software Maintenance & Evolution
A7.2.1. - Decision procedures
A7.2.2. - Automated Theorem Proving
A7.2.3. - Interactive Theorem Proving
A7.3.1. - Computational models and calculability
A9.8. - Reasoning

Other Research Topics and Application Domains:

B9.5.1. - Computer science
B9.5.2. - Mathematics
B9.8. - Reproducibility

1. Team, Visitors, External Collaborators

Research Scientists

Dale Miller [Team leader, Inria, Senior Researcher]
Beniamino Accattoli [Inria, Researcher]
Kaustuv Chaudhuri [Inria, Researcher]
François Lamarche [Inria, Senior Researcher]
Stéphane Graham-Lengrand [CNRS, Researcher]
Gabriel Scherer [Inria, Researcher]
Lutz Straßburger [Inria, Researcher, HDR]

External Collaborator

Andrea Condoluci [University of Bologna (Italy), from Jun 2018 until Aug 2018]

PhD Students

Andrea Condoluci [University of Bologna (Italy), until May 2018]
Ulysse Gerard [Inria]
Maico Carlos Leberle [Inria]
Matteo Manighetti [Inria]
François Thiré [Ecole Normale Supérieure Cachan]

Post-Doctoral Fellow

Matteo Acclavio [Inria, until Nov 2018]

Visiting Scientists

Gopalan Nadathur [University of Minnesota (USA), from Oct 2018 until Nov 2018]
Carlos Olarte [Federal University of Rio Grande do Norte (Brazil), from Oct 2018 until Nov 2018]
Elaine Pimentel [Federal University of Rio Grande do Norte (Brazil), from Oct 2018 until Nov 2018]

Administrative Assistant

Maeva Jeannot [Inria]

2. Overall Objectives

2.1. Main themes

The aim of the Parsifal team is to develop and exploit *proof theory* and *type theory* in the specification, verification, and analysis of computational systems.

- *Expertise*: the team conducts basic research in proof theory and type theory. In particular, the team is developing results that help with automated deduction and with the manipulation and communication of formal proofs.
- *Design*: based on experience with computational systems and theoretical results, the team develops new logical principles, new proof systems, and new theorem proving environments.
- *Implementation*: the team builds prototype systems to help validate basic research results.
- *Examples*: the design and implementation efforts are guided by examples of specification and verification problems. These examples not only test the success of the tools but also drive investigations into new principles and new areas of proof theory and type theory.

The foundational work of the team focuses on *structural* and *analytic* proof theory, *i.e.*, the study of formal proofs as algebraic and combinatorial structures and the study of proof systems as deductive and computational formalisms. The main focus in recent years has been the study of the *sequent calculus* and of the *deep inference* formalisms.

An important research question is how to reason about computational specifications that are written in a *relational* style. To this end, the team has been developing new approaches to dealing with induction, co-induction, and generic quantification. A second important question is of *canonicity* in deductive systems, *i.e.*, when are two derivations "essentially the same"? This crucial question is important not only for proof search, because it gives an insight into the structure and an ability to manipulate the proof search space, but also for the communication of *proof objects* between different reasoning agents such as automated theorem provers and proof checkers.

Important application areas currently include:

- Meta-theoretic reasoning on functional programs, such as terms in the λ-calculus
- Reasoning about behaviors in systems with concurrency and communication, such as the π-calculus, game semantics, *etc.*
- Combining interactive and automated reasoning methods for induction and co-induction
- Verification of distributed, reactive, and real-time algorithms that are often specified using modal and temporal logics
- Representing proofs as documents that can be printed, communicated, and checked by a wide range of computational logic systems.
- Development of cost models for the evaluation of proofs and programs.

3. Research Program

3.1. General overview

There are two broad approaches for computational specifications. In the *computation as model* approach, computations are encoded as mathematical structures containing nodes, transitions, and state. Logic is used to *describe* these structures, that is, the computations are used as models for logical expressions. Intensional operators, such as the modals of temporal and dynamic logics or the triples of Hoare logic, are often employed to express propositions about the change in state.

The *computation as deduction* approach, in contrast, expresses computations logically, using formulas, terms, types, and proofs as computational elements. Unlike the model approach, general logical apparatus such as cut-elimination or automated deduction becomes directly applicable as tools for defining, analyzing, and animating computations. Indeed, we can identify two main aspects of logical specifications that have been very fruitful:

- *Proof normalization*, which treats the state of a computation as a proof term and computation as normalization of the proof terms. General reduction principles such as β-reduction or cut-elimination are merely particular forms of proof normalization. Functional programming is based on normalization [57], and normalization in different logics can justify the design of new and different functional programming languages [30].
- *Proof search*, which views the state of a computation as a a structured collection of formulas, known as a *sequent*, and proof search in a suitable sequent calculus as encoding the dynamics of the computation. Logic programming is based on proof search [61], and different proof search strategies can be used to justify the design of new and different logic programming languages [60].

While the distinction between these two aspects is somewhat informal, it helps to identify and classify different concerns that arise in computational semantics. For instance, confluence and termination of reductions are crucial considerations for normalization, while unification and strategies are important for search. A key challenge of computational logic is to find means of uniting or reorganizing these apparently disjoint concerns.

An important organizational principle is structural proof theory, that is, the study of proofs as syntactic, algebraic and combinatorial objects. Formal proofs often have equivalences in their syntactic representations, leading to an important research question about *canonicity* in proofs – when are two proofs "essentially the same?" The syntactic equivalences can be used to derive normal forms for proofs that illuminate not only the proofs of a given formula, but also its entire proof search space. The celebrated *focusing* theorem of Andreoli [32] identifies one such normal form for derivations in the sequent calculus that has many important consequences both for search and for computation. The combinatorial structure of proofs can be further explored with the use of *deep inference*; in particular, deep inference allows access to simple and manifestly correct cut-elimination procedures with precise complexity bounds.

Type theory is another important organizational principle, but most popular type systems are generally designed for either search or for normalization. To give some examples, the Coq system [70] that implements the Calculus of Inductive Constructions (CIC) is designed to facilitate the expression of computational features of proofs directly as executable functional programs, but general proof search techniques for Coq are rather primitive. In contrast, the Twelf system [66] that is based on the LF type theory (a subsystem of the CIC), is based on relational specifications in canonical form (*i.e.*, without redexes) for which there are sophisticated automated reasoning systems such as meta-theoretic analysis tools, logic programming engines, and inductive theorem provers. In recent years, there has been a push towards combining search and normalization in the same type-theoretic framework. The Beluga system [67], for example, is an extension of the LF type theory with a purely computational meta-framework where operations on inductively defined LF objects can be expressed as functional programs.

The Parsifal team investigates both the search and the normalization aspects of computational specifications using the concepts, results, and insights from proof theory and type theory.

3.2. Inductive and co-inductive reasoning

The team has spent a number of years in designing a strong new logic that can be used to reason (inductively and co-inductively) on syntactic expressions containing bindings. This work is based on earlier work by McDowell, Miller, and Tiu [59] [58] [62] [71], and on more recent work by Gacek, Miller, and Nadathur [44] [43]. The Parsifal team, along with our colleagues in Minneapolis, Canberra, Singapore, and Cachan, have been building two tools that exploit the novel features of this logic. These two systems are the following.

- Abella, which is an interactive theorem prover for the full logic.
- Bedwyr, which is a model checker for the "finite" part of the logic.

We have used these systems to provide formalize reasoning of a number of complex formal systems, ranging from programming languages to the λ-calculus and π-calculus.

Since 2014, the Abella system has been extended with a number of new features. A number of new significant examples have been implemented in Abella and an extensive tutorial for it has been written [1].

3.3. Developing a foundational approach to defining proof evidence

The team is developing a framework for defining the semantics of proof evidence. With this framework, implementers of theorem provers can output proof evidence in a format of their choice: they will only need to be able to formally define that evidence's semantics. With such semantics provided, proof checkers can then check alleged proofs for correctness. Thus, anyone who needs to trust proofs from various provers can put their energies into designing trustworthy checkers that can execute the semantic specification.

In order to provide our framework with the flexibility that this ambitious plan requires, we have based our design on the most recent advances within the theory of proofs. For a number of years, various team members have been contributing to the design and theory of *focused proof systems* [33] [35] [37] [38] [46] [55] [56] and we have adopted such proof systems as the corner stone for our framework.

We have also been working for a number of years on the implementation of computational logic systems, involving, for example, both unification and backtracking search. As a result, we are also building an early and reference implementation of our semantic definitions.

3.4. Deep inference

Deep inference [48], [50] is a novel methodology for presenting deductive systems. Unlike traditional formalisms like the sequent calculus, it allows rewriting of formulas deep inside arbitrary contexts. The new freedom for designing inference rules creates a richer proof theory. For example, for systems using deep inference, we have a greater variety of normal forms for proofs than in sequent calculus or natural deduction systems. Another advantage of deep inference systems is the close relationship to category-theoretic proof theory. Due to the deep inference design one can directly read off the morphism from the derivations. There is no need for a counter-intuitive translation.

The following research problems are investigated by members of the Parsifal team:

- Find deep inference system for richer logics. This is necessary for making the proof theoretic results of deep inference accessible to applications as they are described in the previous sections of this report.
- Investigate the possibility of focusing proofs in deep inference. As described before, focusing is a way to reduce the non-determinism in proof search. However, it is well investigated only for the sequent calculus. In order to apply deep inference in proof search, we need to develop a theory of focusing for deep inference.

3.5. Proof nets, atomic flows, and combinatorial proofs

Proof nets graph-like presentations of sequent calculus proofs such that all "trivial rule permutations" are quotiented away. Ideally the notion of proof net should be independent from any syntactic formalism, but most notions of proof nets proposed in the past were formulated in terms of their relation to the sequent calculus. Consequently we could observe features like “boxes” and explicit “contraction links”. The latter appeared not only in Girard’s proof nets [45] for linear logic but also in Robinson’s proof nets [68] for classical logic. In this kind of proof nets every link in the net corresponds to a rule application in the sequent calculus.

Only recently, due to the rise of deep inference, new kinds of proof nets have been introduced that take the formula trees of the conclusions and add additional “flow-graph” information (see e.g., [54][2] leading to the notion of *atomic flow* and [49]. On one side, this gives new insights in the essence of proofs and their normalization. But on the other side, all the known correctness criteria are no longer available.

Combinatorial proofs [52] are another form syntax-independent proof presentation which separates the multiplicative from the additive behaviour of classical connectives.

The following research questions investigated by members of the Parsifal team:

- Finding (for classical and intuitionistic logic) a notion of canonical proof presentation that is deductive, i.e., can effectively be used for doing proof search.
- Studying the normalization of proofs using atomic flows and combinatorial proofs, as they simplify the normalization procedure for proofs in deep inference, and additionally allow to get new insights in the complexity of the normalization.
- Studying the size of proofs in the combinatorial proof formalism.

3.6. Cost Models and Abstract Machines for Functional Programs

In the *proof normalization* approach, computation is usually reformulated as the evaluation of functional programs, expressed as terms in a variation over the λ-calculus. Thanks to its higher-order nature, this approach provides very concise and abstract specifications. Its strength is however also its weakness: the abstraction from physical machines is pushed to a level where it is no longer clear how to measure the complexity of an algorithm.

Models like Turing machines or RAM rely on atomic computational steps and thus admit quite obvious cost models for time and space. The λ-calculus instead relies on a single non-atomic operation, β-reduction, for which costs in terms of time and space are far from evident.

Nonetheless, it turns out that the number of β-steps is a reasonable time cost model, i.e.,it is polynomially related to those of Turing machines and RAM. For the special case of *weak evaluation* (i.e., reducing only β-steps that are not under abstractions)—which is used to model functional programming languages—this is a relatively old result due to Blelloch and Greiner [34] (1995). It is only very recently (2014) that the strong case—used in the implementation models of proof assistants—has been solved by Accattoli and Dal Lago [31].

With the recent recruitment of Accattoli, the team’s research has expanded in this direction. The topics under investigations are:

1. *Complexity of Abstract Machines*. Bounding and comparing the overhead of different abstract machines for different evaluation schemas (weak/strong call-by-name/value/need λ-calculi) with respect to the cost model. The aim is the development of a complexity-aware theory of the implementation of functional programs.
2. *Reasonable Space Cost Models*. Essentially nothing is known about reasonable space cost models. It is known, however, that environment-based execution model—which are the mainstream technology for functional programs—do not provide an answer. We are exploring the use of the non-standard implementation models provided by Girard’s Geometry of Interaction to address this question.

4. Application Domains

4.1. Trustworthy implementations of theorem proving techniques

The production of real-world verified software has made it necessary to integrate results coming from different theorem provers in a single certification package. One approach to this integration task is by exchanging proof evidence and relying on a backend proof-checker.

Another approach to integration consists in re-implementing the theorem proving techniques as proof-search strategies, on an architecture that guarantees correctness.

Inference systems in general, and focused sequent calculi in particular, can serve as the basis of such an architecture, providing primitives for the exploration of the search space. These form a trusted *Application Programming Interface* that can be used to program and experiment various proof-search heuristics without worrying about correctness. No proof-checking is needed if one trusts the implementation of the API.

This approach has led to the development of the Psyche engine, and to its latest branch CDSAT.

Three major research directions are currently being explored, based on the above:

- The first one is about formulating automated reasoning techniques in terms of inference systems, so that they fit the approach described above. While this is rather standard for technique used in first-order Automated Theorem Provers (ATP), such as resolution, superposition, etc, this is much less standard in SMT-solving, the branch of automated reasoning that can natively handle reasoning in a combination of mathematical theories: the traditional techniques developed there usually organise the collaborations between different reasoning black boxes, whose opaque mechanisms less clearly connect to proof-theoretical inference systems. We are therefore investigating new foundations for reasoning in combinations of theories, expressed as fine-grained inference systems, and developed the *Conflict-Driven Satisfiability framework* for these foundations [13].
- The second one is about understanding how to deal with quantifiers in presence of one or more theories: On the one hand, traditional techniques for quantified problems, such as *unification* [29] or *quantifier elimination* are usually designed for either the empty theory or very specific theories. On the other hand, the industrial techniques for combining theories (Nelson-Oppen, Shostak, MCSAT [64], [69], [73], [53]) are designed for quantifier-free problems, and quantifiers there are dealt with incomplete *clause instantiation* methods or *trigger*-based techniques [41]. We are working on making the two approaches compatible.
- The above architecture's modular approach raises the question of how its different modules can safely cooperate (in terms of guaranteed correctness), while some of them are trusted and others are not. The issue is particularly acute if some of the techniques are run concurrently and exchange data at unpredictable times. For this we explore new solutions based on Milner's *LCF* [63]. In [47], we argued that our solutions in particular provide a way to fulfil the "Strategy Challenge for SMT-solving" set by De Moura and Passmore [74].

4.2. Principled computation for strong lambda-calculi

The application domain of the *cost models and abstract machines for functional programs* line of work—when *application* is intended in concrete terms—is the implementation of proof assistants.

Both functional languages and proof assistants rely on the λ-calculus has reference model. Functional languages are built on the *weak*λ-calculus (where evaluation does not enter function bodies) whose theory is simple and whose implementation has been widely explored in the last decades. Proof assistants instead require the full power of the *strong*λ-calculus, whose theory is more involved and whose implementation has mostly been neglected by the literature.

The study of reasonable cost models naturally leads to a refined theory of implementations, where different techniques and optimisations are classified depending on their complexity (with respect to the cost model). This direction is particularly relevant for the strong λ-calculus, for which most implementations are developed in a *ad-hoc* way.

The theoretical study in particular pointed out that all available proof assistants are implemented following unreasonable implementation schemas, where *unreasonable* here means with potentially exponential overhead with respect to the number of steps in the calculus.

Beniamino Accattoli collaborates with Bruno Barras—one of the implementors of *Coq*, the most used proof assistant—and Claudio Sacerdoti Coen—one of the implementors of *Matita*—in order to develop a fine theory of implementation for proof assistants.

If *applications* are intended also at a more theoretical level, the study of reasonable cost models is also applicable to the development of quantitative denotational semantics, to higher-order approaches to complexity theory, and to implicit computational complexity.

5. Highlights of the Year

5.1. Highlights of the Year

D. Miller has been made General Chair of the LICS Conference Series for three years, starting July 2018.

6. New Software and Platforms

6.1. Abella

FUNCTIONAL DESCRIPTION: Abella is an interactive theorem prover for reasoning about computations given as relational specifications. Abella is particularly well suited for reasoning about binding constructs.

- Participants: Dale Miller, Gopalan Nadathur, Kaustuv Chaudhuri, Mary Southern, Matteo Cimini, Olivier Savary-Bélanger and Yuting Wang
- Partner: Department of Computer Science and Engineering, University of Minnesota
- Contact: Kaustuv Chaudhuri
- URL: http://abella-prover.org/

6.2. Bedwyr

Bedwyr - A proof search approach to model checking
KEYWORD: Model Checker
FUNCTIONAL DESCRIPTION: Bedwyr is a generalization of logic programming that allows model checking directly on syntactic expressions that possibly contain bindings. This system, written in OCaml, is a direct implementation of two recent advances in the theory of proof search.

It is possible to capture both finite success and finite failure in a sequent calculus. Proof search in such a proof system can capture both may and must behavior in operational semantics. Higher-order abstract syntax is directly supported using term-level lambda-binders, the nabla quantifier, higher-order pattern unification, and explicit substitutions. These features allow reasoning directly on expressions containing bound variables.

The distributed system comes with several example applications, including the finite pi-calculus (operational semantics, bisimulation, trace analyses, and modal logics), the spi-calculus (operational semantics), value-passing CCS, the lambda-calculus, winning strategies for games, and various other model checking problems.

- Participants: Dale Miller, Quentin Heath and Roberto Blanco Martinez
- Contact: Dale Miller
- URL: http://slimmer.gforge.inria.fr/bedwyr/

6.3. Checkers

Checkers - A proof verifier
KEYWORDS: Proof - Certification - Verification
FUNCTIONAL DESCRIPTION: Checkers is a tool in Lambda-prolog for the certification of proofs. Checkers consists of a kernel which is based on LKF and is based on the notion of ProofCert.

- Participants: Giselle Machado Nogueira Reis, Marco Volpe and Tomer Libal
- Contact: Tomer Libal
- URL: https://github.com/proofcert/checkers

6.4. Psyche

Proof-Search factorY for Collaborative HEuristics
FUNCTIONAL DESCRIPTION: Psyche is a modular platform for automated or interactive theorem proving, programmed in OCaml and built on an architecture (similar to LCF) where a trusted kernel interacts with plugins. The kernel offers an API of proof-search primitives, and plugins are programmed on top of the API to implement search strategies. This architecture is set up for pure logical reasoning as well as for theory-specific reasoning, for various theories.

RELEASE FUNCTIONAL DESCRIPTION: It is now equipped with the machinery to handle quantifiers and quantifier-handling techniques. Concretely, it uses meta-variables to delay the instantiation of existential variables, and constraints on meta-variables are propagated through the various branches of the search-space, in a way that allows local backtracking. The kernel, of about 800 l.o.c., is purely functional.

- Participants: Assia Mahboubi, Jean-Marc Notin and Stéphane Graham-Lengrand
- Contact: Stéphane Graham-Lengrand
- URL: http://www.lix.polytechnique.fr/~lengrand/Psyche/

6.5. Maetning

FUNCTIONAL DESCRIPTION: Mætning is an automated theorem prover for intuitionistic predicate logic that is designed to disprove non-theorems.

- Contact: Kaustuv Chaudhuri
- URL: https://github.com/chaudhuri/maetning/

6.6. OCaml

KEYWORDS: Functional programming - Static typing - Compilation
FUNCTIONAL DESCRIPTION: The OCaml language is a functional programming language that combines safety with expressiveness through the use of a precise and flexible type system with automatic type inference. The OCaml system is a comprehensive implementation of this language, featuring two compilers (a bytecode compiler, for fast prototyping and interactive use, and a native-code compiler producing efficient machine code for x86, ARM, PowerPC and System Z), a debugger, a documentation generator, a compilation manager, a package manager, and many libraries contributed by the user community.

- Participants: Damien Doligez, Xavier Leroy, Fabrice Le Fessant, Luc Maranget, Gabriel Scherer, Alain Frisch, Jacques Garrigue, Marc Shinwell, Jeremy Yallop and Leo White
- Contact: Damien Doligez
- URL: https://ocaml.org/

7. New Results

7.1. Functional programming with λ-tree syntax

Participants: Ulysse Gerard, Dale Miller, Gabriel Scherer.

We have been designing a new functional programming language, MLTS, that uses the λ-*tree* syntax approach to encoding bindings that appear within data structures [17]. In this setting, bindings never become free nor escape their scope: instead, binders in data structures are permitted to *move* into binders within programs phrases. The design of MLTS—whose concrete syntax is based on that of OCaml—includes additional sites within programs that directly support this movement of bindings. Our description of MLTS includes a typing discipline that naturally extends the typing of OCaml programs.

The operational semantics of MLTS is given using natural semantics for evaluation. We shall view such natural semantics as a logical theory with a rich logic that includes both nominal abstraction and the ∇-quantifier: as a result, the natural semantic specification of MLTS can be given a succinct and elegant presentation.

We have developed a number of examples of how this new programming language can be used. Some of the most convincing of these examples are programs that manipuate untyped λ-terms. A web-based implementation of an MLTS interpreter is available to anyone with a modern web browser: simply visit https://trymlts.github.io/. Small MLTS programs can be composed and executed using that interpreter.

7.2. Proof theory for model checking

Participant: Dale Miller.

While model checking has often been considered as a practical alternative to building formal proofs, we have argued that the theory of sequent calculus proofs can be used to provide an appealing foundation for model checking [7]. Given that the emphasis of model checking is on establishing the truth of a property in a model, our framework concentrates on *additive* inference rules since these provide a natural description of truth values via inference rules. Unfortunately, using these rules alone can force the use of inference rules with an infinite number of premises. In order to accommodate more expressive and finitary inference rules, *multiplicative* rules must be used, but limited to the construction of *additive synthetic inference rules*: such synthetic rules are described using the proof-theoretic notions of polarization and focused proof systems. This framework provides a natural, proof-theoretic treatment of reachability and non-reachability problems, as well as tabled deduction, bisimulation, and winning strategies. (Q. Heath collaborated on several parts of this research effort.)

7.3. From syntactic proofs to combinatorial proofs

Participants: Matteo Acclavio, Lutz Straßburger.

We continued our research on combinatorial proofs as a notion of proof identity for classical logic. We managed to extend our results from last year: We show for various syntactic formalisms including sequent calculus, analytic tableaux, and resolution, how they can be translated into combinatorial proofs, and which notion of identity they enforce. This allows the comparison of proofs that are given in different formalisms.

These results have been presented at the MLA workshop ins Kanazawa and the IJCAR conference in Oxford, published in [25].

7.4. Proof nets for first-order additive linear logic

Participant: Lutz Straßburger.

In a joint work with Willem Heijltjes (University of Bath) and Dominic Hughes (UC Berkeley) we present canonical proof nets for first-order additive linear logic, the fragment of linear logic with sum, product, and first-order universal and existential quantification. We present two versions of our proof nets. One, witness nets, retains explicit witnessing information to existential quantification. For the other, unification nets, this information is absent but can be reconstructed through unification. Unification nets embody a central contribution of the paper: first-order witness information can be left implicit, and reconstructed as needed. Witness nets are canonical for first-order additive sequent calculus. Unification nets in addition factor out any inessential choice for existential witnesses. Both notions of proof net are defined through coalescence, an additive counterpart to multiplicative contractibility, and for witness nets an additional geometric correctness criterion is provided. Both capture sequent calculus cut-elimination as a one-step global composition operation.

These results are published in [26] and have been presented at the First workshop of the Proof Society in Ghent and at the 3rd FISP workshop in Vienna.

7.5. On the Decision Problem for MELL

Participant: Lutz Straßburger.

The decision problem for multiplicative exponential linear logic (MELL) is one of the most important open problems in the are of linear logic. in 2015 there has been an attempt by Bimbò to prove the decidability of MELL. However, we have found several mistakes in that work, and the main mistake is so serious that there is no obvious fix, and therefore the decidability of MELL remains to be open. As a side effect, our work contains a complete (syntactic) proof of the decidability of the relevant version of MELL, that is the logic obtained from MELL by replacing the linear logic contraction rule by a general unrestricted version of the contraction rule. These results are presented in [27].

7.6. OCaml metatheory

Participant: Gabriel Scherer.

We worked on the evolution of advanced features of the OCaml programming language, designing static analyses to ensure their safety through a scientific study their metatheory. Specifically, we worked on unboxed type declarations (during an internship by Simon Colin, M1 from École Polytechnique) and recursive value definitions (during an internship by Alban Reynaud, L3 from ENS Lyon). The two internships and followup work each resulted in both a change proposal to the OCaml implementation and a submission to an academic conference.

7.7. Merlin: understanding a language server

Participant: Gabriel Scherer.

Thomas Réfis (Jane Street) and Frédéric Bour maintain the Merlin language server of OCaml, a tool that provides language-aware features to text editors. We collaborated with them on dissecting the tool and explaining its design and evolution ([4]); the similarities and differences with usual compiler frontends may inform future language implementation work, and our language-agnostic presentation may be of use to tool designers for other languages and proof assistants.

7.8. Language interoperability: ML and a Linear language

Participant: Gabriel Scherer.

In a programming system where programs are created in one programming language, we consider the addition of another programming language that interoperates with the first – and the reimplementation of some library/system functions in this new language. This can increase expressivity, but it could also break some assumptions made by programmers. Typically, adding a bridge to C or assembly code can introduce memory-unsafe code in a previously-safe system. In [18], we formalize a notion of “graceful” interoperability between two languages in this setting, determined by full abstraction, that is, preservation of equational reasoning. We instantiate this general idea by extending ML with an advanced expert language with linear types and linear mutable cells.

7.9. First-class simultaneous substitutions in the two-level logic approach

Participant: Kaustuv Chaudhuri.

The *two-level logic approach* that underlies the Abella prover is excellent at reasoning about the inductive structure of terms with binding constructs, such as λ-terms from the λ-calculus. However, there is no built in support in Abella for reasoning about the inductive structure of (simultaneous) substitutions. This lack of this kind of support is often criticized in the λ-tree syntax representational style that is used in Abella; indeed, in a number of other systems based on this style, support for reasoning about substitutions is explicitly added into the trusted kernel. In [14] we show how to formalize substitutions in Abella in a fluent and high level manner, where all the meta-theory can be proven in a straightforward manner. We illustrate its use in giving a clean formulation of fact that the Howe extension of applicative similarity is a pre-congruence, a standard result from the meta-theory of the λ-calculus that requires sophistication in treating simultaneous substitutions.

7.10. Hybrid Linear Logic, revisited

Participant: Kaustuv Chaudhuri.

Hybrid Linear Logic (HyLL) was proposed by Chaudhuri and Despeyroux in 2010 as a meta-logic for reasoning about constrained transition systems, with applications to a number of domains including formal molecular biology [36]. This logic is an extension of (intuitionistic) linear logic with hybrid connectives that can reason about monoidal constraint domains such as instants of time or rate functions. *Linear logic with subexponential* is a different extension of linear logic that has been proposed as a mechanism for capturing certain well known constrained settings such as bigraphs [39] or concurrent constraint programming [65]. In a paper accepted to MSCS [5] we show how to relate these two extensions of linear logic by giving an embedding of HyLL into linear logic with subexponentials. Furthermore, we show that subexponentials are able to give an adequate encoding of CTL$*$, which is beyond the expressive power of HyLL. Thus, subexponentials appear to be the better choice as a foundation for constraints in linear logic.

7.11. Proof Nets and the Linear Substitution Calculus

Participant: Beniamino Accattoli.

This work [21] belongs to line of work *Cost Models and Abstract Machines for Functional Programs*, supported by the ANR project COCA HOLA, and it has been published in the proceedings of the international conference ICTAC 2018.

The *Linear Substitution Calculus* (LSC) is a refinement of the λ-calculus that is crucial for the study of cost models for functional programs, as it enables a sharp and yet simple decomposition of the evaluation of λ-terms, and it is employed in the proof of various results about cost models in the literature.

In this work we show that the LSC is isomorphic to the linear logic representation of the λ-calculus. More precisely, it is isomorphic to the *proof nets* presentation of such a fragment of linear logic. Proof nets are a graphical formalism, which—as most graphical formalisms—is handy for intuitions but not prone to formal reasoning. The result is relevant because it allows to manipulate formally a graphical formalism (proof nets) by means of an ordinary term syntax (the LSC).

7.12. Tight Typings and Split Bounds

Participants: Beniamino Accattoli, Stéphane Graham-Lengrand.

This joint work with Delia Kesner (Paris Diderot University) [12] belongs to line of work *Cost Models and Abstract Machines for Functional Programs*, supported by the ANR project COCA HOLA, and it has been published in the proceedings of the international conference ICFP 2018.

Intersection types are a classic tool in the study of the λ-calculus. They are known to characterise various termination properties.

It is also well-known that *multi types*, a variant of intersection types strongly related to linear logic, also characterise termination properties. Typing derivation of multi types, moreover, provide quantitative information such as the number of evaluation step and the size of the results, as first shown by de Carvalho.

In this work we provide some new results on this line of work, notably we provide the first quantitative study via multi types of the leftmost and linear head evaluation strategies. Moreover, we show that our approach covers also the other cases in the literature.

7.13. Types of Fireballs

Participant: Beniamino Accattoli.

This joint work with Giulio Guerrieri (Bologna University) [22] belongs to line of work *Cost Models and Abstract Machines for Functional Programs*, supported by the ANR project COCA HOLA, and it has been published in the proceedings of the international conference APLAS 2018.

The theory of the call-by-value λ-calculus has mostly been developed for *closed* programs, that is, programs without free variables. In the last few years, the authors dedicated considerable efforts to extend it to open terms, that is the case relevant for the implementation of proof assistants. The simplest presentation of the call-by-value λ-calculus for open terms is the *fireball calculus*.

In this work we extend the quantitative study via multi types mentioned in *Tight Typings and Split Bounds* to the fireball calculus.

7.14. Decision procedures for intuitionistic propositional logic

Participant: Stéphane Graham-Lengrand.

Provability in intuitionistic propositional logic is decidable and, as revealed by the works of, e.g., Vorobev [72], Hudelmaier [51] and Dyckhoff [42], proof theory can provide natural decision procedures, which have been implemented in various software. More precisely, a decision procedure is obtained by performing direct root-first proof-search in (different variants of) a sequent calculus system called LJT (aka G4ip); termination is ensured by a property of the sequent calculus called depth-boundedness.

Independently from this, Claessen and Rosen [40] recently proposed a decision procedure for the same logic, based on a methodology used in the field of Satisfiability-Modulo-Theories (SMT). Their implementation clearly outperforms the sequent-calculus-based implementations.

In 2018 we managed to establish of formal connection between the G4ip sequent calculus and the algorithm from [40], revealing the features that they share and the features that distinguish them. This connection is interesting because it gives a proof-theoretical light on SMT-solving techniques, and it opens the door to the design of an intuitionistic version of the CDCL algorithm used in SAT-solvers, which decides provability in classical logic.

7.15. Admissible Tools in the Kitchen of Intuitionistic Logic

Participants: Matteo Manighetti, Andrea Condoluci.

In this work we study the computational meaning of the inference rules that are admissible, but not derivable, in intuitionistic logic [16].

An inference rule is admissible for a logic if whenever its antecedent is derivable, its conclusion was already derivable without the rule. In classical logic, whenever this is the case, then also the implication between antecedent and conclusion is derivable. The notion of an admissible rule is therefore internalized in the logic.

This is not the case for intuitionistic logic, and some rules that are admissible are not derivable: therefore they need reasoning outside the usual intuitionistic logic in order to be reduced to purely intuitionistic derivation.

In this work we propose a proof system with term annotations and reduction rules to give a computational meaning to these reductions.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

COCA HOLA: Cost Models for Complexity Analyses of Higher-Order Languages, coordinated by B. Accattoli, 2016–2019.

FISP: The Fine Structure of Formal Proof Systems and their Computational Interpretations, coordinated by Lutz Straßburger in collaboration with Université Paris 7, Universität Innsbruck and TU Wien, 2016–2019.

8.1.2. Competitivity Clusters

UPScale: Universality of Proofs in SaCLay, a Working Group of LabEx DigiCosme, organized by Chantal Keller (LRI) with regular participation from Parsifal members and a post-doc co-supervision.

8.2. International Research Visitors

8.2.1. Internships

Simon Colin did an M1 internship supervised by G. Scherer, conducting a static analysis to check the safety, in OCaml, of unboxing annotations on type declarations.

Alban Reynaud did an L3 internship supervised by G. Scherer, conducting a static analysis to check the safety, in OCaml, of recursive value declarations.

8.2.2. Visits to International Teams

8.2.2.1. Research Stays Abroad

S. Graham-Lengrand was an International Fellow at SRI International, for 25 months over a period of three years between 2015 and 2018.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

D. Miller is the General Chair of LICS (Logic In Computer Science), starting July 2018.

9.1.1.2. Member of the Organizing Committees

D. Miller is on the Steering Committee for the FSCD conference series and the CPP conference series.

D. Miller is a member of the SIGLOG advisory board, starting November 2015.

9.1.2. Scientific Events Selection

9.1.2.1. Chair of Conference Program Committees

B. Accattoli co-chaired LSFA 2018: 13th Workshop on Logical and Semantic Frameworks with Applications, Fortaleza, Brazil, September 26-28, 2018.

G. Scherer chaired ML2018: the ML Family Workshop 2018 in Saint Louis, US, on Friday September 28th 2018.

L. Straßburger chaired TYDI 2018: Workshop on "Twenty Years of Deep Inference" in Oxford July 7, 2018.

9.1.2.2. Member of the Conference Program Committees

B. Accattoli was on the PPDP 2018 Program Committe: 20th International Symposium on Principles and Practice of Declarative Programming, Frankfurt, Germany, 3–5 September 2018.

S. Graham-Lengrand was on the LFMTP 2018 Program Comittee: Workshop on Logical Frameworks and Meta-Languages: Theory and Practice, Oxford, UK, 7 July 2018.

L. Straßburger was on the PC for LACompLing 2018: Symposium on Logic and Algorithms in Computational Linguistics, Stockholm, 28–31 August 2018

D. Miller was on the program committee for IJCAR-2018: 9th International Joint Conference on Automated Reasoning, Oxford, 14-17 July 2018.

D. Miller was a member of the jury for selecting the 2018 Ackermann Award (the EACSL award for outstanding doctoral dissertation in the field of Logic in Computer Science).

Member of the EATCS Distinguished Dissertation Award Committee since March 2013.

G. Scherer was on the POPL 2019 Program Committee: Principles Of Programming Languages, 13-19 January 2019 Cascais/Lisbon, Portugal

9.1.2.3. Reviewer

G. Scherer reviewed for Computer Science Logic (CSL).

L. Straßburger was reviewer for the following conferences:

- LICS 2018
- IJCAR 2018
- FSCD 2018
- AiML 2018
- ARQNL 2018

B. Accattoli reviewed for LICS 2018, FSCD 2018, PPDP 2018, LSFA 2018.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

D. Miller is on the editorial board of the following journals:

- Journal of Automated Reasoning
- Journal of Applied Logics

9.1.3.2. Reviewer - Reviewing Activities

G. Scherer reviewed for Mathematical Structures in Computer Science (MSCS).

L. Straßburger was reviewer for the following journals:

- Transactions on Computational Logic, ToCL (2x)
- Logical Methods in Computer Science, LMCS
- Mathematical Structures in Computer Science, MSCS
- Journal of Logic, Language and Information, JLLI
- Journal of Automated Reasoning, JAR
- Notre Dame Journal of Formal Logic, NDJFL

B. Accattoli reviewed for Logical Methods in Computer Science (LMCS) and Theoretical Computer Science (TCS).

9.1.4. Invited Talks

S. Graham-Lengrand gave an invited talk at the JFLA 2018 (January), and an invited lecture series at the 8th Summer School on Formal Techniques (May).

B. Accattoli gave an invited talk at the *IFIP Working Group 1.6: Rewriting* on July 8 2018 in Oxford, Uk.

D. Miller was an invited speaker and panelist at the Workshop on Proof Theory and its Applications, 6–7 September 2018 in Ghent, Beligum.

D. Miller gave a colloquim talk at the Technical University of Vienna on 31 October 2018 and at the Cyber Security Lab, NTU, Singapore, 21 March 2018.

9.1.5. Scientific Expertise

G. Scherer participated to a scientific expertise of the implementation of the Tezos blockchain – implemented in OCaml.

9.1.6. Research Administration

L. Straßburger was reviewer for the NWO (Netherlands Organisation for Scientific Research).

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence : G. Scherer, Programmation Fonctionnelle, 50, L1, Paris 8 (Vincennes / Saint Denis), France

Licence : K. Chaudhuri, Programmation avancée en OCaml, 40 hours eq TD, L3, École polytechnique, France

Bachelor : K. Chaudhuri, Computer programming, principal instructor, École polytechnique, France (This program has no direct equivalent in the traditional French university system; the closest would be L1.)

Licence: S. Graham-Lengrand, "*INF412: Fondements de l'Informatique: Logique, Modèles, Calcul*", 32 hours eq. TD, L3, École Polytechnique, France.

Master: S. Graham-Lengrand, "*INF551: Computational Logic*", 45 hours eq. TD, M1, École Polytechnique, France.

Master: B. Accattoli, "*Logique linéaire et paradigmes logiques du calcul*", 18 hours eq. TD, M2, Master Parisien de Recherche en Informatique (MPRI), France.

Master: D. Miller, "*Logique linéaire et paradigmes logiques du calcul*", 18 hours eq. TD, M2, Master Parisien de Recherche en Informatique (MPRI), France.

Summer School: B. Accattoli, "The Complexity of Beta-reduction", 4.5h, International School on Rewriting (ISR) 2018, Cali, Colombia.

9.2.2. Supervision

PhD : Sonia Marin, Modal Proof Theory through a Focused Telescope, Université Paris-Saclay, 30 January 2018, encadrant(s): Lutz Straßburger, Dale Miller.

PhD in progress: Ulysse Gérard and Matteo Manighetti supervised by Dale Miller.

PhD in progress: François Thiré (since 1st October 2016), supervised by S. Graham-Lengrand (joint with G. Dowek).

PhD in progress: Maico Leberle supervised by Dale Miller and Beniamino Accattoli.

9.2.3. Juries

D. Miller was the a reporter for the PhD juries of Michael Lettmann (TU Vienna, 30 October 2018)

9.3. Popularization

L. Straßburger serves as member of the "commission développement technologique (CDT)" for Inria Saclay–Île-de-France (since June 2012).

F. Lamarche was site co-ordinator for the Activity Report for Inria Saclay–Ile-de-France.

9.3.1. Interventions

G. Scherer and M. Manighetti participated the “Fête de la Science” exhibit at Inria Saclay on the whole day of October 11th, 2018. They manned an activity on sorting algorithms for colored plastic pieces.

9.3.2. Internal action

G. Scherer spoke at the “Unithé ou café” meeting, a Saclay-internal popularization meeting, on February 1st, 2018.

10. Bibliography

Major publications by the team in recent years

[1] D. BAELDE, K. CHAUDHURI, A. GACEK, D. MILLER, G. NADATHUR, A. TIU, Y. WANG.*Abella: A System for Reasoning about Relational Specifications*, in "Journal of Formalized Reasoning", 2014, vol. 7, n^{o} 2, p. 1-89 [*DOI :* 10.6092/ISSN.1972-5787/4650], https://hal.inria.fr/hal-01102709

[2] A. GUGLIELMI, T. GUNDERSEN, L. STRASSBURGER.*Breaking Paths in Atomic Flows for Classical Logic*, in "Proceedings of the 25th Annual IEEE Symposium on Logic in Computer Science (LICS 2010)", Edinburgh, United Kingdom, July 2010, p. 284–293 [*DOI :* 10.1109/LICS.2010.12], http://www.lix.polytechnique.fr/~lutz/papers/AFII.pdf

Publications of the year

Articles in International Peer-Reviewed Journal

[3] M. ACCLAVIO.*Proof Diagrams for Multiplicative Linear Logic: Syntax and Semantics*, in "Journal of Automated Reasoning", May 2018 [*DOI :* 10.1007/S10817-018-9466-4], https://hal.archives-ouvertes.fr/hal-01931400

[4] F. BOUR, T. RÉFIS, G. SCHERER.*Merlin: a language server for OCaml (experience report)*, in "Proceedings of the ACM on Programming Languages", July 2018, vol. 2, n^{o} ICFP, p. 1 - 15, https://arxiv.org/abs/1807.06702 [*DOI :* 10.1145/3236798], https://hal.inria.fr/hal-01929161

[5] K. CHAUDHURI, C. OLARTE, E. PIMENTEL, J. DESPEYROUX.*Hybrid Linear Logic, revisited*, in "Mathematical Structures in Computer Science", 2019, https://hal.inria.fr/hal-01968154

[6] O. FLÜCKIGER, G. SCHERER, M.-H. YEE, A. GOEL, A. AHMED, J. VITEK.*Correctness of Speculative Optimizations with Dynamic Deoptimization*, in "Proceedings of the ACM on Programming Languages", 2018, vol. 2, n^{o} POPL, https://arxiv.org/abs/1711.03050 [*DOI :* 10.1145/3158137], https://hal.inria.fr/hal-01646765

[7] Q. HEATH, D. MILLER.*A proof theory for model checking*, in "Journal of Automated Reasoning", 2018 [*DOI :* 10.1007/S10817-018-9475-3], https://hal.inria.fr/hal-01814006

[8] J. KOPPEL, G. SCHERER, A. SOLAR-LEZAMA.*Capturing the Future by Replaying the Past Functional Pearl*, in "Proceedings of the ACM on Programming Languages", July 2018, vol. 2, n^{o} ICFP, p. 1 - 29 [*DOI :* 10.1145/3236771], https://hal.inria.fr/hal-01929178

[9] R. KUZNETS, L. STRASSBURGER.*Maehara-style modal nested calculi*, in "Archive for Mathematical Logic", July 2018 [*DOI :* 10.1007/s00153-018-0636-1], https://hal.inria.fr/hal-01942240

[10] D. MILLER.*Mechanized metatheory revisited*, in "Journal of Automated Reasoning", 2018, https://hal.inria.fr/hal-01884210

[11] L. STRASSBURGER.*Deep inference and expansion trees for second-order multiplicative linear logic*, in "Mathematical Structures in Computer Science", November 2018, p. 1-31 [*DOI :* 10.1017/S0960129518000385], https://hal.inria.fr/hal-01942410

International Conferences with Proceedings

[12] B. ACCATTOLI, S. GRAHAM-LENGRAND, D. KESNER.*Tight typings and split bounds*, in "23rd ACM International Conference on Functional Programming", St Louis, United States, M. FLATT (editor), September 2018, vol. 2, n^{o} ICFP, p. 1 - 30, https://arxiv.org/abs/1807.02358 [*DOI :* 10.1145/3236789], https://hal.archives-ouvertes.fr/hal-01936141

[13] M. P. BONACINA, S. GRAHAM-LENGRAND, N. SHANKAR.*Proofs in conflict-driven theory combination*, in "Proceedings of the 7th International Conference on Certified Programs and Proofs (CPP'18)", Los Angeles, United States, J. ANDRONICK, A. FELTY (editors), ACM Press, January 2018, vol. 18 [*DOI :* 10.1145/3167096], https://hal.archives-ouvertes.fr/hal-01935595

[14] K. CHAUDHURI.*A Two-Level Logic Perspective on (Simultaneous) Substitutions*, in "CPP 2018 - Proceedings of the 7th ACM SIGPLAN International Conference on Certified Programs and Proofs", Los Angeles, United States, January 2018, https://hal.inria.fr/hal-01968139

[15] K. CHAUDHURI, U. GÉRARD, D. MILLER.*Computation-as-deduction in Abella: work in progress*, in "13th international Workshop on Logical Frameworks and Meta-Languages: Theory and Practice", Oxford, United Kingdom, July 2018, https://hal.inria.fr/hal-01806154

[16] A. CONDOLUCI, M. MANIGHETTI.*Admissible Tools in the Kitchen of Intuitionistic Logic*, in "Seventh International Workshop on Classical Logic and Computation (CL&C 2018)", Oxford, United Kingdom, Electronic Proceedings in Theoretical Computer Science, July 2018, vol. 281, p. 10-23, https://hal.inria.fr/hal-01870112

[17] U. GÉRARD, D. MILLER.*Functional programming with $\lambda-tree$ syntax: a progress report*, in "13th international Workshop on Logical Frameworks and Meta-Languages: Theory and Practice", Oxford, United Kingdom, July 2018, https://hal.inria.fr/hal-01806129

[18] G. SCHERER, M. NEW, N. RIOUX, A. AHMED.*FabULous Interoperability for ML and a Linear Language*, in "International Conference on Foundations of Software Science and Computation Structures (FoSSaCS)", Thessaloniki, Greece, C. BAIER, U. D. LAGO (editors), FabOpen image in new windowous Interoperability for ML and a Linear Language, Springer, April 2018, vol. LNCS - Lecture Notes in Computer Science, n^{o} 10803, https://arxiv.org/abs/1707.04984 [*DOI :* 10.1007/978-3-319-89366-2_8], https://hal.inria.fr/hal-01929158

National Conferences with Proceeding

[19] G. BARANY, G. SCHERER.*Génération aléatoire de programmes guidée par la vivacité*, in "JFLA 2018 - Journées Francophones des Langages Applicatifs", Banyuls-sur-Mer, France, January 2018, https://hal.inria.fr/hal-01682691

[20] S. COLIN, R. LEPIGRE, G. SCHERER.*Unboxing Mutually Recursive Type Definitions in OCaml*, in "JFLA 2019", Les Rousses, France, January 2019, https://arxiv.org/abs/1811.02300 , https://hal.inria.fr/hal-01929508

Conferences without Proceedings

[21] B. ACCATTOLI.*Proof Nets and the Linear Substitution Calculus*, in "15th International Colloquium on Theoretical Aspects of Computing (ICTAC 2018)", Stellenbosch, South Africa, October 2018, https://hal.archives-ouvertes.fr/hal-01967532

[22] B. ACCATTOLI, G. GUERRIERI.*Types of Fireballs*, in "16th Asian Symposium on Programming Languages and System (APLAS 2018)", Wellington, New Zealand, December 2018, https://hal.archives-ouvertes.fr/hal-01967531

[23] F. BOBOT, S. GRAHAM-LENGRAND, B. MARRE, G. BURY.*Centralizing equality reasoning in MCSAT*, in "16th International Workshop on Satisfiability Modulo Theories (SMT 2018)", Oxford, United Kingdom, R. DIMITROVA, V. D'SILVA (editors), July 2018, https://hal.archives-ouvertes.fr/hal-01935591

[24] S. GRAHAM-LENGRAND, M. FÄRBER.*Guiding SMT solvers with Monte Carlo Tree Search and neural networks*, in "Third Conference on Artificial Intelligence and Theorem Proving (AITP'2018)", Aussois, France, T. C. HALES, C. KALISZYK, S. SCHULZ, J. URBAN (editors), March 2018, https://hal.archives-ouvertes.fr/hal-01935578

Scientific Books (or Scientific Book chapters)

[25] M. ACCLAVIO, L. STRASSBURGER.*From Syntactic Proofs to Combinatorial Proofs*, in "International Joint Conference on Automated Reasoning, IJCAR 2018", Springer, June 2018, p. 481-497, https://hal.inria.fr/hal-01942275

Research Reports

[26] W. HEIJLTJES, D. J. D. HUGHES, L. STRASSBURGER.*Proof nets for first-order additive linear logic*, Inria, September 2018, n^{o} RR-9201, https://hal.inria.fr/hal-01867625

[27] L. STRASSBURGER.*On the Decision Problem for MELL*, Inria Saclay Ile de France, September 2018, n^{o} 9203, https://hal.inria.fr/hal-01870148

Other Publications

[28] D. MILLER.*Influences between logic programming and proof theory*, March 2018, HaPoP 2018: Fourth Symposium on the History and Philosophy of Programming, https://hal.inria.fr/hal-01900891

References in notes

[29] J. A. ROBINSON, A. VORONKOV (editors). *Handbook of Automated Reasoning*, Elsevier and MIT press, 2001

[30] S. ABRAMSKY.*Computational Interpretations of Linear Logic*, in "Theoretical Computer Science", 1993, vol. 111, p. 3–57

[31] B. ACCATTOLI, U. DAL LAGO.*Beta reduction is invariant, indeed*, in "Joint Meeting of the Twenty-Third EACSL Annual Conference on Computer Science Logic (CSL) and the Twenty-Ninth Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), CSL-LICS '14, Vienna, Austria, July 14 - 18, 2014", 2014, p. 8:1–8:10, http://doi.acm.org/10.1145/2603088.2603105

[32] J.-M. ANDREOLI.*Logic Programming with Focusing Proofs in Linear Logic*, in "Journal of Logic and Computation", 1992, vol. 2, n^o 3, p. 297–347

[33] D. BAELDE, D. MILLER, Z. SNOW.*Focused Inductive Theorem Proving*, in "Fifth International Joint Conference on Automated Reasoning (IJCAR 2010)", J. GIESL, R. HÄHNLE (editors), LNCS, 2010, n^o 6173, p. 278–292 [*DOI :* 10.1007/978-3-642-14203-1], http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/ijcar10.pdf

[34] G. E. BLELLOCH, J. GREINER.*Parallelism in Sequential Functional Languages*, in "Proceedings of the seventh international conference on Functional programming languages and computer architecture, FPCA 1995, La Jolla, California, USA, June 25-28, 1995", 1995, p. 226–237, http://doi.acm.org/10.1145/224164.224210

[35] K. CHAUDHURI.*The Focused Inverse Method for Linear Logic*, Carnegie Mellon University, December 2006, Technical report CMU-CS-06-162, http://reports-archive.adm.cs.cmu.edu/anon/2006/CMU-CS-06-162.pdf

[36] K. CHAUDHURI, J. DESPEYROUX.*A Hybrid Linear Logic for Constrained Transition Systems with Applications to Molecular Biology*, Inria, 2010

[37] K. CHAUDHURI, N. GUENOT, L. STRASSBURGER.*The Focused Calculus of Structures*, in "Computer Science Logic: 20th Annual Conference of the EACSL", Leibniz International Proceedings in Informatics (LIPIcs), Schloss Dagstuhl–Leibniz-Zentrum für Informatik, September 2011, p. 159–173 [*DOI :* 10.4230/LIPIcs.CSL.2011.159], http://drops.dagstuhl.de/opus/volltexte/2011/3229/pdf/16.pdf

[38] K. CHAUDHURI, S. HETZL, D. MILLER.*A Multi-Focused Proof System Isomorphic to Expansion Proofs*, in "Journal of Logic and Computation", June 2014 [*DOI :* 10.1093/LOGCOM/EXU030], http://hal.inria.fr/hal-00937056

[39] K. CHAUDHURI, G. REIS.*An adequate compositional encoding of bigraph structure in linear logic with subexponentials*, in "Logic for Programming, Artificial Intelligence, and Reasoning (LPAR)", M. DAVIS, A. FEHNKER, A. MCIVER, A. VORONKOV (editors), LNCS, Springer, November 2015, vol. 9450, p. 146–161 [*DOI :* 10.1007/978-3-662-48899-7_11], http://chaudhuri.info/papers/lpar15bigraphs.pdf

[40] K. CLAESSEN, D. ROSÉN.*SAT Modulo Intuitionistic Implications*, in "Proc. of the the 20th Int. Conf. on Logic for Programming, Artificial Intelligence, and Reasoning (LPAR'15)", M. DAVIS, A. FEHNKER, A. MCIVER, A. VORONKOV (editors), LNCS, Springer-Verlag, November 2015, vol. 9450, p. 622–637, http://dx.doi.org/10.1007/978-3-662-48899-7

[41] C. DROSS, S. CONCHON, J. KANIG, A. PASKEVICH.*Reasoning with Triggers*, in "10th Intern. Worksh. on Satisfiability Modulo Theories, SMT 2012", P. FONTAINE, A. GOEL (editors), EPiC Series, EasyChair, June 2012, vol. 20, p. 22–31, http://smt2012.loria.fr/paper4.pdf

[42] R. DYCKHOFF. *Contraction-free sequent calculi for intuitionistic logic*, in "J. of Symbolic Logic", 1992, vol. 57, n^o 3, p. 795–807

[43] A. GACEK, D. MILLER, G. NADATHUR. *Combining generic judgments with recursive definitions*, in "23th Symp. on Logic in Computer Science", F. PFENNING (editor), IEEE Computer Society Press, 2008, p. 33–44, http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/lics08a.pdf

[44] A. GACEK, D. MILLER, G. NADATHUR. *Nominal abstraction*, in "Information and Computation", 2011, vol. 209, n^o 1, p. 48–73, http://arxiv.org/abs/0908.1390

[45] J.-Y. GIRARD. *Linear Logic*, in "Theoretical Computer Science", 1987, vol. 50, p. 1–102

[46] S. GRAHAM-LENGRAND, R. DYCKHOFF, J. MCKINNA. *A Focused Sequent Calculus Framework for Proof Search in Pure Type Systems*, in "Logical Methods in Computer Science", 2011, vol. 7, n^o 1, http://www.lix.polytechnique.fr/~lengrand/Work/Reports/TTSC09.pdf

[47] S. GRAHAM-LENGRAND. *Slot Machines: an approach to the Strategy Challenge in SMT solving (presentation only)*, in "13th International Workshop on Satisfiability Modulo Theories", San Francisco, United States, July 2015, https://hal.inria.fr/hal-01211209

[48] A. GUGLIELMI. *A System of Interaction and Structure*, in "ACM Trans. on Computational Logic", 2007, vol. 8, n^o 1

[49] A. GUGLIELMI, T. GUNDERSEN. *Normalisation Control in Deep Inference Via Atomic Flows*, in "Logical Methods in Computer Science", 2008, vol. 4, n^o 1:9, p. 1–36, http://arxiv.org/abs/0709.1205

[50] A. GUGLIELMI, L. STRASSBURGER. *Non-commutativity and MELL in the Calculus of Structures*, in "Computer Science Logic, CSL 2001", L. FRIBOURG (editor), LNCS, Springer-Verlag, 2001, vol. 2142, p. 54–68

[51] J. HUDELMAIER. *Bounds for Cut Elimination in Intuitionistic Logic*, Universität Tübingen, 1989

[52] D. HUGHES. *Proofs Without Syntax*, in "Annals of Mathematics", 2006, vol. 164, n^o 3, p. 1065–1076

[53] D. JOVANOVIĆ, C. BARRETT, L. DE MOURA. *The Design and Implementation of the Model Constructing Satisfiability Calculus*, in "Proc. of the 13th Int. Conf. on Formal Methods In Computer-Aided Design (FMCAD '13)", FMCAD Inc., 2013, Portland, Oregon, http://www.cs.nyu.edu/~barrett/pubs/JBdM13.pdf

[54] F. LAMARCHE, L. STRASSBURGER. *Naming Proofs in Classical Propositional Logic*, in "Typed Lambda Calculi and Applications, TLCA 2005", P. URZYCZYN (editor), LNCS, Springer, 2005, vol. 3461, p. 246–261

[55] C. LIANG, D. MILLER. *Focusing and Polarization in Linear, Intuitionistic, and Classical Logics*, in "Theoretical Computer Science", 2009, vol. 410, n^o 46, p. 4747–4768, http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/tcs09.pdf

[56] C. LIANG, D. MILLER. *A Focused Approach to Combining Logics*, in "Annals of Pure and Applied Logic", 2011, vol. 162, n^o 9, p. 679–697, http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/lku.pdf

[57] P. MARTIN-LÖF.*Constructive Mathematics and Computer Programming*, in "Sixth International Congress for Logic, Methodology, and Philosophy of Science", Amsterdam, North-Holland, 1982, p. 153–175

[58] R. MCDOWELL, D. MILLER.*Reasoning with Higher-Order Abstract Syntax in a Logical Framework*, in "ACM Trans. on Computational Logic", 2002, vol. 3, n^o 1, p. 80–136, http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/mcdowell01.pdf

[59] R. MCDOWELL, D. MILLER.*A Logic for Reasoning with Higher-Order Abstract Syntax*, in "Proceedings, Twelfth Annual IEEE Symposium on Logic in Computer Science", Warsaw, Poland, G. WINSKEL (editor), IEEE Computer Society Press, July 1997, p. 434–445

[60] D. MILLER.*Forum: A Multiple-Conclusion Specification Logic*, in "Theoretical Computer Science", September 1996, vol. 165, n^o 1, p. 201–232

[61] D. MILLER, G. NADATHUR, F. PFENNING, A. SCEDROV.*Uniform Proofs as a Foundation for Logic Programming*, in "Annals of Pure and Applied Logic", 1991, vol. 51, p. 125–157

[62] D. MILLER, A. TIU.*A Proof Theory for Generic Judgments: An extended abstract*, in "Proc. 18th IEEE Symposium on Logic in Computer Science (LICS 2003)", IEEE, June 2003, p. 118–127, http://www.lix.polytechnique.fr/Labo/Dale.Miller/papers/lics03.pdf

[63] R. MILNER.*LCF: A Way of Doing Proofs with a Machine*, in "Proc. of the 8th Intern. Symp. on the Mathematical Foundations of Computer Science", J. BECVÁR (editor), LNCS, Springer, 1979, vol. 74, p. 146-159

[64] G. NELSON, D. C. OPPEN.*Simplification by Cooperating Decision Procedures*, in "ACM Press Trans. on Program. Lang. and Syst.", October 1979, vol. 1, n^o 2, p. 245–257, http://dx.doi.org/10.1145/357073.357079

[65] C. OLARTE, V. NIGAM, E. PIMENTEL.*Subexponential concurrent constraint programming*, in "Theoretical Computer Science", 2015, http://dx.doi.org/10.1016/j.tcs.2015.06.031

[66] F. PFENNING, C. SCHÜRMANN.*System Description: Twelf — A Meta-Logical Framework for Deductive Systems*, in "16th Conference on Automated Deduction", Trento, H. GANZINGER (editor), LNAI, Springer, 1999, n^o 1632, p. 202–206

[67] B. PIENTKA, J. DUNFIELD.*Beluga: A Framework for Programming and Reasoning with Deductive Systems (System Description)*, in "Fifth International Joint Conference on Automated Reasoning", J. GIESL, R. HÄHNLE (editors), LNCS, 2010, n^o 6173, p. 15–21

[68] E. P. ROBINSON.*Proof Nets for Classical Logic*, in "Journal of Logic and Computation", 2003, vol. 13, p. 777–797

[69] R. E. SHOSTAK.*Deciding Combinations of Theories*, in "J. ACM", 1984, vol. 31, n^o 1, p. 1–12, http://dx.doi.org/10.1145/2422.322411

[70] THE COQ DEVELOPMENT TEAM.*The Coq Proof Assistant Version 8.3 Reference Manual*, Inria, October 2010

[71] A. TIU, D. MILLER.*Proof Search Specifications of Bisimulation and Modal Logics for the π-calculus*, in "ACM Trans. on Computational Logic", 2010, vol. 11, n^{o} 2, http://arxiv.org/abs/0805.2785

[72] N. N. VOROB'EV.*A new algorithm for derivability in the constructive propositional calculus*, in "Amer. Math. Soc. Transl.", 1970, vol. 94, n^{o} 2, p. 37–71

[73] L. M. DE MOURA, D. JOVANOVIĆ.*A Model-Constructing Satisfiability Calculus*, in "Proc. of the 14th Int. Conf. on Verification, Model Checking, and Abstract Interpretation (VMCAI'13)", R. GIACOBAZZI, J. BERDINE, I. MASTROENI (editors), LNCS, Springer-Verlag, 2013, vol. 7737, p. 1–12, http://dx.doi.org/10.1007/978-3-642-35873-9_1

[74] L. M. DE MOURA, G. O. PASSMORE.*The Strategy Challenge in SMT Solving*, in "Automated Reasoning and Mathematics - Essays in Memory of William W. McCune", M. P. BONACINA, M. E. STICKEL (editors), LNCS, Springer, 2013, vol. 7788, p. 15–44 [*DOI :* 10.1007/978-3-642-36675-8_2], http://dx.doi.org/10.1007/978-3-642-36675-8

Project-Team PETRUS

PErsonal & TRUSted cloud

IN PARTNERSHIP WITH:
Université Versailles Saint-Quentin

RESEARCH CENTER
Saclay - Île-de-France

THEME
Data and Knowledge Representation and Processing

Table of contents

Project-Team PETRUS

Creation of the Team: 2016 December 01, updated into Project-Team: 2017 July 01

Keywords:

Computer Science and Digital Science:
- A1.1.8. - Security of architectures
- A1.4. - Ubiquitous Systems
- A3.1.2. - Data management, quering and storage
- A3.1.3. - Distributed data
- A3.1.5. - Control access, privacy
- A3.1.6. - Query optimization
- A3.1.8. - Big data (production, storage, transfer)
- A3.1.9. - Database
- A4.3. - Cryptography
- A4.5. - Formal methods for security
- A4.7. - Access control
- A4.8. - Privacy-enhancing technologies

Other Research Topics and Application Domains:
- B2.5.3. - Assistance for elderly
- B6.4. - Internet of things
- B6.6. - Embedded systems
- B9.10. - Privacy

1. Team, Visitors, External Collaborators

Research Scientists
- Nicolas Anciaux [Team leader, Inria, Senior Researcher, HDR]
- Luc Bouganim [Inria, Senior Researcher, HDR]

Faculty Members
- Philippe Pucheral [Univ de Versailles Saint-Quentin-en-Yvelines, Professor, HDR]
- Iulian Sandu Popa [Univ de Versailles Saint-Quentin-en-Yvelines, Associate Professor]
- Guillaume Scerri [Univ de Versailles Saint-Quentin-en-Yvelines, Associate Professor]

External Collaborator
- Benjamin Nguyen [INSA Centre Val-de-Loire, Professor, HDR]

Technical Staff
- Aydogan Ersoz [Inria]
- Laurent Schneider [Inria]

PhD Students
- Robin Carpentier [Univ de Versailles Saint-Quentin-en-Yvelines]
- Dimitrios Tsolovos [Inria]
- Riad Ladjel [Inria]
- Julien Loudet [CozyCloud]
- Axel Michel [INSA CVL]
- Paul Tran Van [CozyCloud]

Administrative Assistants

Adeline Lochet [Inria, from Jun 2018]

Emmanuelle Perrot [Inria]

2. Overall Objectives

2.1. Overall Objectives

We are witnessing an exponential accumulation of personal data on central servers: data automatically gathered by administrations and companies but also data produced by individuals themselves (e.g., photos, agendas, data produced by smart appliances and quantified-self devices) and deliberately stored in the cloud for convenience. The net effect is, on the one hand, an unprecedented threat on data privacy due to abusive usage and attacks and, on the other hand, difficulties in providing powerful user-centric services (e.g. personal big data) which require crossing data stored today in isolated silos. The Personal Cloud paradigm holds the promise of a Privacy-by-Design storage and computing platform, where each individual can gather her complete digital environment in one place and share it with applications and users, while preserving her control. However, this paradigm leaves the privacy and security issues in user's hands, which leads to a paradox if we consider the weaknesses of individuals' autonomy in terms of computer security, ability and willingness to administer sharing policies. The challenge is however paramount in a society where emerging economic models are all based - directly or indirectly - on exploiting personal data.

While many research works tackle the organization of the user's workspace, the semantic unification of personal information, the personal data analytics problems, the objective of the PETRUS project-team is to tackle the privacy and security challenges from an architectural point of view. More precisely, our objective is to help providing a technical solution to the personal cloud paradox. More precisely, our goals are (i) to propose new architectures (encompassing both software and hardware aspects) for secure personal cloud data management and formally prove important bricks of the architecture, (ii) propose new data administration models reaching the main requirements of a personal cloud (decentralized access and usage control models, data sharing, data collection and retention models, etc.) and study the enforcement of the resulting privacy policies based on secure hardware and formally proven architectural components, (iii) propose new secure distributed database indexing models, privacy preserving query processing strategies and data anonymization techniques for the personal cloud.

3. Research Program

3.1. Research Program

To tackle the challenge introduced above, we identify four main lines of research:

- (Axis 1) Personal cloud server architectures. Based on the intuition that user control, security and privacy are key properties in the definition of trusted personal cloud solutions, our objective is to propose new architectures (encompassing both software and hardware aspects) for secure personal cloud data management and formally prove important bricks of the architecture.
- (Axis 2) Privacy preserving administration models and enforcement. This research axis is devoted to the definition of sharing rules that are easily manageable for the individual and enforced by default (i.e., secure implementation). Complementary to the definition of sharing policies, it is mandatory to help the average user regulate the complete lifecycle of her data, from its capture, to its dissemination and up to its deletion. Our objective is to propose new data administration models reaching the main requirements of a personal cloud (decentralized access and usage control models, data sharing, data collection and retention models, etc.) and study the enforcement of the resulting privacy policies based on secure hardware and formally proven architectural components.

- (Axis 3) Global query evaluation. The goal of this line of research is to provide capabilities for crossing data belonging to multiple individuals (e.g., performing statistical queries over personal data, computing queries on social graphs or organizing participatory data collection) in a fully decentralized setting while providing strong and personalized privacy guarantees. This means proposing new secure distributed database indexing models, privacy preserving query processing strategies and data anonymization techniques for the personal cloud.
- (Axis 4) Economic, legal and societal issues. This research axis is more transversal and entails multidisciplinary research, addressing the links between economic, legal, societal and technological aspects. We will follow here a multi-disciplinary approach based on a 3-step methodology: i) identifying important common issues related to privacy and to the exploitation of personal data; ii) characterizing their dimensions in all relevant disciplines and jointly study their entanglement; iii) validating the proposed analysis, models and trade-offs thanks to in vivo experiments.

These contributions will also rely on tools (algorithms, protocols, proofs, etc.) from other communities, namely security (cryptography, secure multiparty computations, formal methods, differential privacy, etc.) and distributed systems (distributed hash tables, gossip protocols, etc.). Beyond the research actions, we structure our software activity around a single common platform (rather than isolated demonstrators), integrating our main research contributions, called PlugDB. This platform is the cornerstone to help validating our research results through accurate performance measurements on a real platform, a common practice in the DB community, and target the best conferences. It is also a strong vector to federate the team, simplify the bootstrapping of new PhD or master students, conduct multi-disciplinary research and open the way to industrial collaborations and technological transfers.

4. Application Domains

4.1. Personal cloud, home care, IoT, sensing, surveys

As stated in the software section, the Petrus research strategy aims at materializing its scientific contributions in an advanced hardware/software platform with the expectation to produce a real societal impact. Hence, our software activity is structured around a common Secure Personal Cloud platform rather than several isolated demonstrators. This platform will serve as the foundation to develop a few emblematic applications. Several privacy-preserving applications can actually be targeted by a Personal Cloud platform, like: (i) smart disclosure applications allowing the individual to recover her personal data from external sources (e.g., bank, online shopping activity, insurance, etc.), integrate them and cross them to perform personal big data tasks (e.g., to improve her budget management) ; (ii) management of personal medical records for care coordination and well-being improvement; (iii) privacy-aware data management for the IoT (e.g., in sensors, quantified-self devices, smart meters); (iv) community-based sensing and community data sharing; (v) privacy-preserving studies (e.g., cohorts, public surveys, privacy-preserving data publishing). Such applications overlap with all the research axes described above but each of them also presents its own specificities. For instance, the smart disclosure applications will focus primarily on sharing models and enforcement, the IoT applications require to look with priority at the embedded data management and sustainability issues, while community-based sensing and privacy-preserving studies demand to study secure and efficient global query processing. Among these applications domains, one is already receiving a particular attention from our team. Indeed, we gained a strong expertise in the management and protection of healthcare data through our past DMSP (Dossier Medico-Social Partagé) experiment in the field. This expertise is being exploited to develop a dedicated healthcare and well-being personal cloud platform. We are currently deploying 10000 boxes equipped with PlugDB in the context of the DomYcile project. In this context, we are currently setting up an Inria Innovation Lab with the Hippocad company to industrialize this platform and deploy it at large scale (see Section the bilateral contract OwnCare II-Lab).

5. Highlights of the Year

5.1. Highlights of the Year

Creation of the Inria Innovation Lab 'OwnCare'

PETRUS has set up the OwnCare Inria Innovation Lab (IILab) with UVSQ and the Hippocad company in January 2018. The objective of this IILab is to industrialize PlugDB, a flagship software/hardware platform initiated in the SMIS team and today pursued in PETRUS, and deploy it in the medical/social field. A first deployment over 10.000 patients is planned in the Yvelines district (see Section 8.1.1 for details).

6. New Software and Platforms

6.1. PLUG-DB ENGINE

KEYWORDS: Databases - Personal information - Privacy - Hardware and Software Platform
FUNCTIONAL DESCRIPTION: en PlugDB is a complete platform dedicated to a secure and ubiquitous management of personal data. It aims at providing an alternative to a systematic centralization of personal data. The PlugDB engine is a personal database server capable of storing data (tuples and documents) in tables and BLOBs, indexing them, querying them in SQL, sharing them through assertional access control policies and enforcing transactional properties (atomicity, integrity, durability).

The PlugDB engine is embedded in a tamper-resistant hardware device combining the security of smartcard with the storage capacity of NAND Flash. The personal database is hosted encrypted in NAND Flash and the PlugDB engine code runs in the microcontroller. Complementary modules allow to pre-compile SQL queries for the applications, communicate with the DBMS from a remote Java program, synchronize local data with remote servers (typically used for recovering the database in the case of a broken or lost devices) and participate in distributed computation (e.g., global queries). PlugDB runs both on secure devices provided by Gemalto and on specific secure devices designed by PETRUS and assembled by electronic SMEs. Mastering the hardware platform opens up new research and experiment opportunities (e.g., support for wireless communication, secure authentication, sensing capabilities, battery powered ...). PlugDB engine has been registered first at APP (Agence de Protection des Programmes) in 2009 - a new version being registered every two years - and the hardware datasheets in 2015.

PlugDB has been experimented in the field, notably in the healthcare domain. We also recently set up an educational platform on top of PlugDB, named SIPD (Système d'Information Privacy-by-Design) and used at ENSIIE, INSA CVL and UVSQ through the Versailles Sciences Lab fablab, to raise students awareness of privacy protection problems and embedded programming. As a conclusion, PlugDB combines several research contributions from the team, at the crossroads of flash data management, embedded data processing and secure distributed computations. It then strongly federates all members of our team (permanent members, PhD students and engineers). It is also a vector of visibility, technological transfer and dissemination and gives us the opportunity to collaborate with researchers from other disciplines around a concrete privacy-enhancing platform.

PlugDB is now being industrialized in the context of the OwnCare Inria Innovation Lab (II-Lab). In OwnCare, PlugDB acts as a secure personal cloud to manage medical/social data for people receiving care at home. It should be deployed over 10.000 patient in the Yvelines district. The industrialization process covers the development of a complete testing environment, the writing of a detailed documentation and the development of additional features (e.g., embedded ODBC driver, TPM support, flexible access control model and embedded code upgrade notably). It has also required the design of a new hardware platform equipped with a battery power supply, introducing new energy consumption issues for the embedded software.

- Participants: Aydogan Ersoz, Laurent Schneider, Luc Bouganim, Nicolas Anciaux and Philippe Pucheral
- Contact: Nicolas Anciaux
- URL: https://project.inria.fr/plugdb/

7. New Results

7.1. Extensive and Secure PDMS Architecture (Axis 1)

Participants: Nicolas Anciaux [correspondent], Luc Bouganim, Philippe Pucheral, Iulian Sandu Popa, Guillaume Scerri, Dimitrios Tsolovos.

The Personal Cloud paradigm is emerging through a myriad of solutions offered to users to let them gather and manage their whole digital life. This paradigm shift towards user empowerment raises fundamental questions with regards to the appropriateness of the data management functionalities and protection techniques which are offered by existing solutions to laymen users. This year, we reviewed, compared and analyzed personal cloud alternatives in terms of the functionalities they provide and the threat models they target. From this analysis, we derived a general set of security requirements that any Personal Data Management System (PDMS) should consider. We then identified the challenges of implementing such a PDMS and proposed a preliminary design for an extensive and secure PDMS reference architecture satisfying the considered requirements. Finally, we discussed several important research challenges remaining to be addressed to achieve a mature PDMS ecosystem. A first paper making the functionality and security standpoint in PDMS solutions, proposing five security goals and a preliminary architecture to fulfill these goal based on Trusted Execution Environments was published at IS'19 [12], and preliminary results on the case of a crowdsensing architecture was presented at Middleware'18 [15] and BDA'18 [18].

7.2. Data sharing model for the Personal Cloud (Axis 2)

Participants: Nicolas Anciaux [correspondent], Philippe Pucheral, Guillaume Scerri, Paul Tran Van, Baptiste Crepin.

In the PDMS context, new sharing models are needed to help end-users controlling the sharing policies under use. We proposed an architecture to produce authorizations satisfying users' sharing desires without having to trust the underlying producing these authorizations in the PhD thesis of Paul Tran-Van [11] and we demonstrated the solution at EDBT'18 [14]. We currently investigate the case of a data sharing system producing what we call 'zero-knowledge permissions', i.e., a set of authorizations produced by an untrusted sharing model which is supposed to reveal no information at all about a given subset of documents in the user space.

7.3. SEP2P: Secure and Efficient P2P Personal Data Processing (Axis 3)

Participants: Luc Bouganim [correspondent], Julien Loudet, Iulian Sandu Popa.

Personal Data Management Systems (PDMS) arrive at a rapid pace allowing us to integrate all our personal data in a single place and use it for our benefit and for the benefit of the community. This leads to a significant paradigm shift since personal data become massively distributed and opens an important question: how can users/applications execute queries and computations over this massively distributed data in a secure and efficient way, relying exclusively on peer-to-peer (P2P) interactions? We studied the feasibility of such a pure P2P personal data management system and provide efficient and scalable mechanisms to reduce the data leakage to its minimum with covert adversaries. In particular, we showed that data processing tasks can be assigned to nodes in a verifiable random way, which cannot be influenced by malicious colluding nodes. We proposed a generic solution which largely minimizes the verification cost. Our experimental evaluation shows that the proposed protocols lead to minimal private information leakage, while the cost of the security mechanisms remains very low even with a large number of colluding corrupted nodes. We illustrated our generic protocol proposal on three data-oriented use-cases, namely, participatory sensing, targeted data diffusion and more general distributed aggregate queries. The full protocol was simulated and evaluated. A first paper focusing on imposed randomness was published at EDBT'19 [13].

7.4. Mobile Participatory Sensing with Strong Privacy Guarantees (Axis 3)

Participant: Iulian Sandu Popa [correspondent].

Mobile participatory sensing could be used in many applications such as vehicular traffic monitoring, pollution tracking, or even health surveying. However, its success depends on finding a solution for querying large numbers of smart phones or vehicular systems, which protects user location privacy and works in real-time. This work proposes PAMPAS, a privacy-aware mobile distributed system for efficient data aggregation in mobile participatory sensing. In PAMPAS, mobile devices enhanced with secure hardware, called secure probes (SPs), perform distributed query processing, while preventing users from accessing other users' data. A supporting server infrastructure (SSI) coordinates the inter-SP communication and the computation tasks executed on SPs. PAMPAS ensures that SSI cannot link the location reported by SPs to the user identities even if SSI has additional background information. Moreover, we propose an enhanced version of the protocol, named PAMPAS$^+$, to make the system robust even against advanced hardware attacks on the SPs. Hence, the user location privacy leakage remains very low even for an attacker controlling the SSI and a few corrupted SPs. The leakage is proportional with the number of corrupted SPs and thus requires a massive SP corruption to break the system, which is extremely unlikely in practice. This work has been accomplished in collaboration with NJIT (see Section 9.2.1.1) and has been recently submitted as a journal paper.

7.5. Trustworthy Distributed Queries on Personal Data using TEEs (Axis 3)

Participants: Riad Ladjel [correspondent], Nicolas Anciaux, Philippe Pucheral, Guillaume Scerri.

The decentralized way of managing personal data in a PDMS provides a de facto protection against massive attacks usually performed on central servers. But this raises the question of how to preserve individuals' trust on their PDMS when performing global computations crossing data from multiple individuals? And how to guarantee the integrity of the final result when it has been computed by a myriad of collaborative but independent PDMSs? We study a secure decentralized computing framework where each participant gains the assurance that his data is only used for the purpose he consents to and that only the final result is disclosed. Conversely, the goal is to provides the querier with the guarantee that this result has been honesty computed, by the expected code on the expected data. A preliminary solution which capitalizes on the use of Trusted Execution Environments (TEE) at the edge of the network was presented at BDA'18 [19] and APVP'18 [20].

7.6. Performance of large scale data-oriented operations under TEE constraints (Axis 3)

Participants: Robin Carpentier [correspondent], Nicolas Anciaux, Iulian Sandu Popa, Guillaume Scerri.

The rise of Trusted Execution Environments like Intel SGX, and their more and more widespread use for data processing raises the question of their impact on performance, specifically for data oriented operations. While some works aim at embedding either the entirety of part of a database engine within a TEE, the direct impact of processing data with TEEs as opposed to more classical environment has not been studied yet. In particular, the cryptographic overhead of accessing persistent data outside the TEE enclave, the limited RAM amount of each TEE enclave, the cost of external function calls and memory access overheads, may slow the computing by orders of magnitude compared to a regular environment, and have to be taken into account. Preliminary results presenting both a benchmark of data operations within Intel SGX, together with optimisation of search algorithm dealing with the specific way of accessing external memory from inside SGX have been presented at BDA'18 [16].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. OwnCare II-Lab (Jul 2017 - Dec 2020)

Partners: PETRUS (Inria-UVSQ), Hippocad (SME)

End 2016, the Yvelines district lauched a public call for tender to deploy an industrial solution aiming at covering the whole distinct (10.000 patients). The Hippocad company, in partnership with Inria, won this call for tender with a solution called DomYcile in May 2017 and the project was launched in July 2017. DomYcile is based on a home box combining the PlugDB hardware/software technology developed by the Petus team and a communication layer based on SigFox. Hippocad and Petrus then decided to launch a joint II-Lab (Inria Innovation Lab) named OwnCare. The objective is threefold: (1) build an industrial solution based on PlugDB and deploy it in the Yvelines district in the short-term, (2) use this Yvelines testbed to improve the solution and try to deploy it at the national/international level in the medium-term and (3) design flexible/secure/mobile personal medical folder solutions targeting individual uses rather than professional uses in the long-term. The DomYcile project with the Yvelines district has started in July 2017 and the II-Lab was officially created in January 2018.

8.2. Bilateral Grants with Industry

8.2.1. Cozy Cloud CIFRE - Tran Van contract (Oct 2014 -Feb 2018)

Partners: Cozy Cloud, PETRUS
Following a bilateral contract with Cozy Cloud (a French startup providing a personal Cloud platform), the CIFRE PhD thesis of Paul Tran Van capitalized on the Cozy-PlugDB platform to devise new access and usage control models to exchange data among devices of the same user (devices may have different levels of trustworthiness) and among different users thanks to a user-friendly sharing model [14].

8.2.2. Cozy Cloud CIFRE - Loudet contract (Apr 2016 - Apr 2019)

Partners: Cozy Cloud, PETRUS
In relation with the bilateral contract mentioned above, a second CIFRE PhD thesis has been started by Julien Loudet. The objective is to allow for a secure execution of distributed queries on a set of personal clouds associated to users, depending on social links, user's localization or user's profile. The general idea is to build secure indexes, distributed on the users' personal clouds and to devise a secure execution protocol revealing solely the query result to the querier. Such highly distributed secure queries potentially enable new (social) applications fed by user's personal data which could be developed on the Cozy-PlugDB platform.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR PerSoCloud (Jan 2017 - Dec 2020)

Partners: Orange Labs (coordinator), PETRUS (Inria-UVSQ), Cozy Cloud, U. of Versailles.
The objective of PerSoCloud is to design, implement and validate a full-fledged Privacy-by-Design Personal Cloud Sharing Platform. One of the major difficulties linked to the concept of personal cloud lies in organizing and enforcing the security of the data sharing while the data is no longer under the control of a central server. We identify three dimensions to this problem. Devices-sharing: assuming that the primary copy of user U1's personal data is hosted in a secure place, how to share and synchronize it with U1's multiple (mobile) devices without compromising security? Peers-sharing: how user U1 could exchange a subset of his-her data with an identified user U2 while providing to U1 tangible guarantees about the usage made by U2 of this data? Community-sharing: how user U1 could exchange a subset of his-her data with a large community of users and contribute to personal big data analytics while providing to U1 tangible guarantees about the preservation of his-her anonymity? In addition to tackling these three scientific and technical issues, a legal analysis will guarantee compliance of this platform with the security and privacy French and UE regulation, which firmly promotes the Privacy by Design principle, including the current reforms of personal data regulation.

9.1.2. PIA - PDP SECSi (May 2016 - Dec 2017)

Partners: Cozy Cloud (coordinator), Qwant, PETRUS (Inria-UVSQ), FING.
The objective of this PIA-PDP (Programme Investissement d'Avenir - Protection des Données Personnelles) SECSi project is to build a concrete Personal Cloud platform which can support a large scale deployment of Self Data services. Three major difficulties are identified and will be tackled in this project: (1) how to implement and enforce a fine control of the data flow when personal data are exploited by third party applications, (2) how to protect these same applications when processing is delegated to the personal cloud platform itself and (3) how to implement personalized search on the web without hurting user's privacy.

9.1.3. CityLab@Inria, Inria Project Lab (May 2014 - Oct 2018)

Inria Partners: ARLES-MIMOVE, CLIME, DICE, FUN, MYRIADS, OAK, PETRUS, URBANET, WILLOW. External partners: UC Berkeley.
CityLab@Inria studies ICT solutions toward smart cities that promote both social and environmental sustainability. A strong emphasis of the Lab is on the undertaking of a multi-disciplinary research program through the integration of relevant scientific and technology studies, from sensing up to analytics and advanced applications, so as to actually enact the foreseen smart city Systems of Systems. PETRUS contributes to Privacy-by-Design architectures for trusted smart objects so as to ensure privacy to citizens, which is critical for ensuring that urbanscale sensing contributes to social sustainability and does not become a threat. The PhD Thesis of Dimitris Tsoulovos, co-directed by MIMOVE and PETRUS, is funded by CityLab. http://citylab.inria.fr/

9.1.4. GDP-ERE, DATA-IA project (Sept. 2018 - Aug. 2021)

Partners: DANTE (U. of Versailles), PETRUS (Inria-UVSQ).
The role of individuals and the control of their data is a central issue in the new European regulation (GDPR) enforced on 25th May 2018. Data portability is a new right provided under those regulations. It allows citizens to retrieve their personal data from the companies and governmental agencies that collected them, in an interoperable digital format. The goals are to enable the individual to get out of a captive ecosystem, and to favor the development of innovative personal data services beyond the existing monopolistic positions. The consequence of this new right is the design and deployment of technical platforms, commonly known as Personal Cloud. But personal cloud architectures are very diverse, ranging from cloud based solutions where millions of personal cloud are managed centrally, to self-hosting solutions. These diversity is not neutral both in terms of security and from the point of view of the chain of liabilities. The GDP-ERE project tends to study those issues in an interdisciplinary approach by the involvement of jurists and computers scientists. The two main objectives are (i) to analyze the effects of the personal cloud architectures on legal liabilities, enlightened by the analysis of the rules provided under the GDPR and (ii) to propose legal and technological evolutions to highlight the share of liability between each relevant party and create adapted tools to endorse those liabilities. http://dataia.eu/actualites/linstitut-dataia-vous-presente-le-projet-gdp-ere-rgpd-et-cloud-personnel-de-lempowerment

9.2. International Research Visitors

9.2.1. Visits to International Teams

9.2.1.1. Research Stays Abroad

Iulian Sandu Popa has visited the Computer Science department of NJIT (New Jersey Institute of Technology) for two months (Mars to April) during 2018. Iulian has a long history of collaboration with this department at NJIT, this being his second long stay since 2011. In particular, he collaborates at NJIT with Professor Vincent Oria on topics related to spatiotemporal data management and with Professor Cristian Borcea on topics such as privacy-preserving mobile computing for location-based applications [5] and secure and distributed crowd-sensing for smart city applications. For the latter topic, a joint journal paper has been recently submitted (see Section 7.4).

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organization

10.1.1.1. Member of the Organizing Committees

- Luc Bouganim: Co-organizer of Ecole thématique BDA Masses de Données Distribuées, Aussois, June 2018
- Iulian Sandu Popa: 34ème Conférence sur la Gestion de Données – Principes, Technologies et Applications (BDA 2018), Bucarest, 22-26 octobre 2018
- Iulian Sandu Popa: Colloque National Capteurs et Sciences Participatives (CASPA), Paris, 1-4 avril 2019

10.1.2. Scientific Events Selection

10.1.2.1. Member of the Conference Program Committees

- Nicolas Anciaux: VLDB'18, VLDB'19, SIGMOD'19, DATA'18, BDA'18
- Luc Bouganim: Associate Editor for VLDB'18
- Philippe Pucheral: DATA'18, MOBILITY'18
- Iulian Sandu Popa: DATA'18, ICDE'19, IEEE MobileCloud'19

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Nicolas Anciaux: Associate Editor of the VLDB Journal

10.1.3.2. Reviewer - Reviewing Activities

- Iulian Sandu Popa: ACM Transactions on Spatial Algorithms and Systems, IEEE Transactions on Parallel and Distributed Systems, Geoinformatica

10.1.4. Invited Talks

- Iulian Sandu Popa: "Highly Distributed Queries on Personal Data Management Systems with Strong Privacy Guarantees", New Jersey Institute of Technology (NJIT), Newark, April 9, 2018. https://web.njit.edu/cs/CS_Seminar/abstract.php?id=391
- Célia Zolynski and Nicolas Anciaux: "L'avènement d'une gestion individuelle de nos données personnelles. Regards croisés sur certains enjeux de vie privée", Kick-Off of the DATAIA Institute, Fev. 2018. slides: http://petrus.inria.fr/~anciaux/papers/D36.pdf
- Philippe Pucheral: "Privacy-by-Obligation: Une réponse au paradoxe du Cloud Personnel ?", Inria Scientific days, June 28th, 2018.
- Célia Zolynski and Nicolas Anciaux: "GDPR and Personal Cloud: from Empowerment to Responsibility (GDP-ERE)", DATAIA-JST International Symposium on Data Science and AI, July 2018. Link: http://dataia.eu/actualites/dataia-jst-international-symposium-data-science-and-ai, slides: http://dataia.eu/sites/default/files/DATAIA_JST_International_Symposium/DATAIA-JST_SYMPOSIUM_Nicolas_Anciaux_Celia_Zolynski.pdf
- Philippe Pucheral: "Protection de la vie privée : potentiel et paradoxe du Cloud Personnel", SystemX seminar, October 18th, 2018. https://www.irt-systemx.fr/philippe-pucheral-animera-un-seminarsystemx-le-18-octobre/
- Nicolas Anciaux: "Personal Data Management Systems using Trusted Execution Environments", Zenith seminar, Inria / Univ. Montpellier, Nov. 2018. https://team.inria.fr/zenith/zenith-seminar-nicolas-anciaux-personal-data-management-systems-using-trusted-execution-environments-21-nov-2018/

10.1.5. Research Administration

- Philippe Pucheral: Member of the HDR committee of the STV doctoral school (UVSQ) since 2014
- Philippe Pucheral: Member of the steering committee of the ED STIC doctoral school of University Paris-Saclay, 'Data, Knowledge and Interactions' committee (about 250 PhD students) since 2014
- Philippe Pucheral: Member of the bureau of the DAVID lab board since 2016
- Nicolas Anciaux: Member of the Council of the Doctoral College of the University Paris-Saclay
- Nicolas Anciaux: Correspondent for the Doctoral school ED STIC of University Paris-Saclay at Inria Saclay
- Nicolas Anciaux: Responsible for the 'Mission Jeunes Chercheurs' (MJC) at Inria Saclay
- Nicolas Anciaux: Responsible for the 'Formation par la Recherche' (FPR) at Inria Saclay
- Nicolas Anciaux: Member of the bureau of the DAVID lab board
- Luc Bouganim: Member of the Scientific Commission (CS) of Inria Saclay-IDF (Cordi-S, Post-Doc, Delegation)
- Luc Bouganim: Member of the Commission for Technological Development (CDT) of Inria Saclay-IDF

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- Licence : Iulian Sandu Popa, Bases de données (niveau L3), 96, UVSQ, France. Guillaume Scerri, Initiation aux bases de données (niveau L2), 63, UVSQ, France. Guillaume Scerri, Fondements de l'informatique (niveau L1), 36, UVSQ, France. Guillaume Scerri, Théorie des Langages (niveau L2), 45, UVSQ, France.
- Master : Iulian Sandu Popa, Bases de données relationnelles (niveau M1), Gestion des données spatiotemporelles (niveau M2), Sécurité des bases de données (niveau M2), 96, UVSQ, France. Philippe Pucheral, responsible of the DataScale master, courses in M1 and M2 in databases and in security, introductory courses for jurists,UVSQ, France. Luc Bouganim, Bases de données relationnelles et XML (niveau M1 et M2), 40, AFTI, France. Guillaume Scerri, Bases de données relationnelles (niveau M1), 36, UVSQ, France. Guillaume Scerri, Sécurité et bases de données pour juristes, 4.5, UVSQ, France. Guillaume Scerri, Sécurité, 18, UVSQ, France.
- Engineers school : Nicolas Anciaux, courses on Databases (module IN206, niveau M1), 21, and Advanced databases (module ASI13, niveau M2), 24, at ENSTA ParisTech. Nicolas Anciaux, Systèmes d'Information "privacy by design" (niveau M1), 30, at ENSIIE Evry, France. Luc Bouganim, Systèmes d'Information "privacy by design" (niveau M1), 42, ENSIIE Evry et INSA CVL, France.

10.2.2. Supervision

- PhD : Paul Tran Van, Partage de documents sécurisé dans le Cloud Personnel, UVSQ, April 3, 2018, Nicolas Anciaux and Philippe Pucheral
- PhD in progress: Axel Michel, Secure Distributed Computations, October 2015, Benjamin Nguyen and Philippe Pucheral
- PhD in progress : Julien Loudet, Highly Distributed Queries on Personal Data Management Systems with Strong Privacy Guarantees, July 2016, Luc Bouganim and Iulian Sandu Popa
- PhD in progress: Riad Ladjel, Secure Distributed Computation for the Personal Cloud, October 2016, Nicolas Anciaux, Philippe Pucheral and Guillaume Scerri
- PhD in progress: Dimitris Tsoulovos, Privacy-by-design Middleware for Urban-scale Mobile Crowdsensing, April 2017, Nicolas Anciaux and Valérie Issarny (Inria Mimove)

- PhD in progress: Robin Carpentier, Secure and efficient data processing in trusted execution environments for the personal cloud, October 2018, Nicolas Anciaux, Iulian Sandu Popa and Guillaume Scerri

10.2.3. Juries

- Nicolas Anciaux : Reviewer of the PhD of Sakina MAHBOUBI (Université de Montpellier, 21/11/2018)
- Luc Bouganim : Reviewer of the PhD of Arezki Laga (University of Bretagne Sud, 20/12/2018)

10.3. Popularization

- Nicolas Anciaux: "Respectez les données personnelles de vos clients avec PLUGDB", Inria Tech Talk, French Tech Central - Station F, 12 Dec. 2018. video: https://youtu.be/9y3VdMe_sAQ, slides: https://french-tech-central.com/events/inria-tech-talk-respectez-les-donnees-personnelles-de-vos-clients-avec-plugdb/
- Nicolas Anciaux: auditions de la mission de préfiguration du « Health Data Hub », DREES, July 2018. Slides: http://petrus.inria.fr/~anciaux/papers/D39.pdf
- Nicolas Anciaux: interview, magazine La Recherche N°535, Mai 2018, "IA : les défis de la stratégie française", pp.14-16, by Gautier Cariou. Link: https://www.larecherche.fr/parution/mensuel-535

11. Bibliography

Major publications by the team in recent years

[1] N. Anciaux, L. Bouganim, P. Pucheral, Y. Guo, L. Le Folgoc, S. Yin.*MILo-DB: a personal, secure and portable database machine*, in "Distributed and Parallel Databases", March 2014, vol. 32, n^o 1, p. 37-63 [*DOI :* 10.1007/s10619-012-7119-x], https://hal.archives-ouvertes.fr/hal-00768355

[2] N. Anciaux, S. Lallali, I. Sandu Popa, P. Pucheral.*A Scalable Search Engine for Mass Storage Smart Objects*, in "41th International Conference on Very Large Databases (VLDB)", Kohala Coast, Hawaii, United States, August 2015, vol. 8, n^o 9, p. 910-921 [*DOI :* 10.14778/2777598.2777600], https://hal.inria.fr/hal-01176458

[3] N. Anciaux, B. Nguyen, I. Sandu Popa.*Tutorial: Managing Personal Data with Strong Privacy Guarantees*, in "17th International Conference on Extending Database Technology (EDBT)", Athens, Greece, March 2014, p. 672-673 [*DOI :* 10.5441/002/edbt.2014.71], https://hal.inria.fr/hal-01096633

[4] S. Lallali, N. Anciaux, I. Sandu Popa, P. Pucheral.*Supporting secure keyword search in the personal cloud*, in "Information Systems", December 2017, vol. 72, p. 1 - 26 [*DOI :* 10.1016/j.is.2017.09.003], https://hal.inria.fr/hal-01660599

[5] S. J. Pan, I. Sandu Popa, C. Borcea.*DIVERT: A Distributed Vehicular Traffic Re-Routing System for Congestion Avoidance*, in "IEEE Transactions on Mobile Computing", January 2017, vol. 16, n^o 1, p. 58-72 [*DOI :* 10.1109/TMC.2016.2538226], https://hal.inria.fr/hal-01426424

[6] G. Scerri, B. Warinschi, M. Barbosa, B. Portela.*Foundations of Hardware-Based Attested Computation and Application to SGX*, in "IEEE European Symposium on Security and Privacy, EuroS&P 2016", Saarbrücken, Germany, March 2016, p. 245-260 [*DOI :* 10.1109/EuroSP.2016.28], https://hal.inria.fr/hal-01417137

[7] C. Q. TO, B. NGUYEN, P. PUCHERAL.*TrustedMR: A Trusted MapReduce System based on Tamper Resistance Hardware* , in "23rd International Conference on Cooperative Information Systems (COOPIS)", Rhodes, Greece, October 2015, p. 38-56 [*DOI :* 10.1007/978-3-319-26148-5_3], https://hal.inria.fr/hal-01254951

[8] C. Q. TO, B. NGUYEN, P. PUCHERAL.*Private and Scalable Execution of SQL Aggregates on a Secure Decentralized Architecture*, in "ACM Transactions on Database Systems", 2016, vol. 41, n^{o} 3, p. 16:1-16:43, https://hal.archives-ouvertes.fr/hal-01296432

[9] D. H. TON THAT, I. SANDU POPA, K. ZEITOUNI.*TRIFL: A Generic Trajectory Index for Flash Storage*, in "ACM Transactions on Algorithms", July 2015, vol. 1, n^{o} 2, 44 [*DOI :* 10.1145/2786758], https://hal.inria.fr/hal-01176563

[10] P. TRAN-VAN, N. ANCIAUX, P. PUCHERAL.*SWYSWYK: A Privacy-by-Design Paradigm for Personal Information Management Systems*, in "International Conference on Information Systems Development (ISD)", Cyprus, Cyprus, September 2017, https://hal.inria.fr/hal-01675090

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] P. TRAN-VAN.*Secure document sharing through Personnal Cloud*, Université Paris-Saclay, April 2018, https://tel.archives-ouvertes.fr/tel-01779315

Articles in International Peer-Reviewed Journal

[12] N. ANCIAUX, P. BONNET, L. BOUGANIM, B. NGUYEN, P. PUCHERAL, I. SANDU POPA, G. SCERRI.*Personal Data Management Systems: The security and functionality standpoint*, in "Information Systems", February 2019, vol. 80, p. 13 - 35 [*DOI :* 10.1016/J.IS.2018.09.002], https://hal.archives-ouvertes.fr/hal-01898705

International Conferences with Proceedings

[13] J. LOUDET, I. SANDU POPA, L. BOUGANIM.*SEP2P: Secure and Efficient P2P Personal Data Processing*, in "EDBT: 22nd International Conference on Extending Database Technology", Lisbon, Portugal, March 2019, https://hal.inria.fr/hal-01949641

[14] P. TRAN-VAN, N. ANCIAUX, P. PUCHERAL.*Reconciling Privacy and Data Sharing in a Smart and Connected Surrounding*, in "International Conference on Extending Database Technology (EDBT)", Vienna, Austria, March 2018, https://hal.inria.fr/hal-01675093

[15] D. TSOLOVOS.*Enforcing Privacy in Participatory Sensing Systems*, in "Middleware Doctoral Symposium 2018", Rennes, France, ACM, December 2018, https://hal.inria.fr/hal-01910067

National Conferences with Proceeding

[16] R. CARPENTIER, N. ANCIAUX, I. S. POPA, G. SCERRI.*Performance of Large Scale Data-Oriented Operations under the TEE Constraints*, in "34ème Conférence sur la Gestion de Données – Principes, Technologies et Applications (BDA 2018)", Bucharest, Romania, October 2018, https://hal.inria.fr/hal-01947896

[17] J. LOUDET, L. BOUGANIM, I. SANDU POPA.*Privacy-Preserving Queries on Highly Distributed Personal Data Management Systems*, in "34ème Conférence sur la Gestion de Données – Principes, Technologies et Applications", Bucharest, Romania, Proceedings of the BDA 2018 Conference, October 2018, https://hal.inria.fr/hal-01949583

[18] D. TSOLOVOS, N. ANCIAUX, V. ISSARNY.*A Privacy Aware Approach for Participatory Sensing Systems*, in "34ème Conférence sur la Gestion de Données – Principes, Technologies et Applications", Bucharest, Romania, October 2018, https://hal.inria.fr/hal-01947863

Conferences without Proceedings

[19] R. LADJEL, N. ANCIAUX, P. PUCHERAL, G. SCERRI.*Secure and Distributed Computations for a Personal Cloud*, in "APVP 2018 – l'Atelier sur la Protection de la Vie Privée", porquerolles, France, June 2018, https://hal.inria.fr/hal-01947842

[20] R. LADJEL, N. ANCIAUX, G. SCERRI, P. PUCHERAL.*Secure and Distributed Computations for a Personal Data Management System*, in "34ème Conférence sur la Gestion de Données – Principes, Technologies et Applications", Bucarest, Romania, October 2018, https://hal.inria.fr/hal-01947808

Scientific Books (or Scientific Book chapters)

[21] A. MICHEL, B. NGUYEN, P. PUCHERAL.*The Case for Personalized Anonymization of Database Query Results*, in "Data Management Technologies and Applications (DATA'17 Revised Selected Papers)", J. FILIPE, J. BERNARDINO, C. QUI (editors), Communications in Computer and Information Science, Springer, June 2018, vol. 814, p. 261-285, https://hal.archives-ouvertes.fr/hal-01945674

Team POEMS-POST

Wave propagation: mathematical analysis and simulation

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

RESEARCH CENTER
Saclay - Île-de-France

THEME
Networks and Telecommunications

Table of contents

Team POEMS-POST

Creation of the Team: 2018 January 01

Keywords:

Computer Science and Digital Science:

A6. - Modeling, simulation and control
A6.1. - Methods in mathematical modeling
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.2. - Stochastic Modeling
A6.1.4. - Multiscale modeling
A6.1.5. - Multiphysics modeling
A6.1.6. - Fractal Modeling
A6.2. - Scientific computing, Numerical Analysis & Optimization
A6.2.1. - Numerical analysis of PDE and ODE
A6.2.2. - Numerical probability
A6.2.3. - Probabilistic methods
A6.2.7. - High performance computing
A6.3.1. - Inverse problems
A6.3.4. - Model reduction
A6.5.1. - Solid mechanics
A6.5.2. - Fluid mechanics
A6.5.4. - Waves

Other Research Topics and Application Domains:

B2.6. - Biological and medical imaging
B3.3. - Geosciences
B3.3.1. - Earth and subsoil
B3.4. - Risks
B3.4.1. - Natural risks
B3.4.2. - Industrial risks and waste
B5.3. - Nanotechnology
B5.4. - Microelectronics
B5.5. - Materials

1. Team, Visitors, External Collaborators

Research Scientists

Marc Bonnet [CNRS, Senior Researcher, HDR]
Anne-Sophie Bonnet-Ben Dhia [CNRS, Team leader, Senior Researcher, HDR]
Eliane Bécache [Inria, Researcher, HDR]
Stéphanie Chaillat [CNRS, Researcher]
Christophe Hazard [CNRS, Researcher, HDR]
Patrick Joly [Inria, Senior Researcher, HDR]
Maryna Kachanovska [Inria, Researcher]

Marc Lenoir [CNRS, Senior Researcher, until Sep 2018]
Jean-François Mercier [CNRS, Senior Researcher, HDR]
Axel Modave [CNRS, Researcher]

Faculty Members

Laurent Bourgeois [École Nationale Supérieure de Techniques Avancées, Professor, HDR]
Patrick Ciarlet [École Nationale Supérieure de Techniques Avancées, Professor, HDR]
Sonia Fliss [École Nationale Supérieure de Techniques Avancées, Professor]
Ennio Fedrizzi [École Nationale Supérieure de Techniques Avancées, Assistant Professor, until Aug 2018]
Laure Giovangigli [École Nationale Supérieure de Techniques Avancées, Assistant Professor, from Oct 2018]
Eric Lunéville [École Nationale Supérieure de Techniques Avancées, Professor]

External Collaborator

Francis Collino [Auto-entrepreneur]

Technical Staff

Colin Chambeyron [CNRS, Research Engineer]
Nicolas Kielbasiewicz [CNRS, Research Engineer]
Christophe Mathulik [École Nationale Supérieure de Techniques Avancées]

PhD Students

Yassine Abourrig [CEA]
Clement Beneteau [École Nationale Supérieure de Techniques Avancées]
Antoine Bensalah [École Nationale Supérieure de Techniques Avancées, until Jul 2018]
Antoine Bera [Ministère de l'Education Nationale]
Damien Chicaud [École Nationale Supérieure de Techniques Avancées, from Oct 2018]
Léandre Giret [CEA, until Jun 2018]
Damien Mavaleix Marchessoux [Naval Group]
Hajer Methenni [CEA]
Sandrine Paolantoni [École Nationale Supérieure de Techniques Avancées]
Emile Parolin [Inria]
Arnaud Recoquillay [CEA, until Jan 2018]
Yohanes Tjandrawidjaja [CEA]

Post-Doctoral Fellows

Faisal Amlani [École Nationale Supérieure de Techniques Avancées]
Felix Kpadonou [Bourse FMJH]
Florian Monteghetti [Inria, from Dec 2018]
Laure Pesudo [École Nationale Supérieure de Techniques Avancées, until May 2018]
Dmitry Ponomarev [École Nationale Supérieure de Techniques Avancées, until Oct 2018]

Administrative Assistants

Natalia Alves [Inria, from Mar 2018]
Corinne Chen [École Nationale Supérieure de Techniques Avancées]

2. Overall Objectives

2.1. The topic of waves

The propagation of waves is one of the most common physical phenomena in nature. From the human scale (sounds, vibrations, water waves, telecommunications, radar) to the scales of the universe (electromagnetic waves, gravity waves) and of the atoms (spontaneous or stimulated emission, interferences between particles), the emission and the reception of waves are our privileged way to understand the world that surrounds us. The study and the simulation of wave propagation phenomena constitute a very broad and active field of research in various domains of physics and engineering sciences. The variety and the complexity of the underlying problems, their scientific and industrial interest, the existence of a common mathematical structure to these problems from different areas altogether justify a research project in applied mathematics and scientific computing devoted to this topic.

2.2. POEMS activities

The project POEMS is an UMR (Unité Mixte de Recherche) between CNRS, ENSTA ParisTech and Inria (UMR 7231). The general activity of the project is oriented toward the design, the analysis, the numerical approximation and the control of mathematical models for the description of wave propagation in mechanics, physics and engineering sciences.

Beyond the general objective of contributing to the progress of the scientific knowledge, four goals can be ascribed to the project:

- the development of expertise relative to various types of waves (acoustic, elastic, electromagnetic, gravity waves, ...), their modelling and numerical simulation,
- the treatment of complex problems whose simulation is close enough to real life situations and industrial applications,
- the development of original mathematical and numerical techniques,
- the development of computational codes, in particular in collaboration with external partners (scientists from other disciplines, industry, state companies...)

3. Research Program

3.1. General description

Our activity relies on the existence of boundary value problems established by physicists to model the propagation of waves in various situations. The basic ingredient is a partial differential equation of the hyperbolic type, whose prototype is the wave equation (or the Helmholtz equation if time-periodic solutions are considered). Nowadays, the numerical techniques for solving the basic academic problems are well mastered. However, the solution of complex wave propagation problems close to real applications still raises (essentially open) problems which constitute a real challenge for applied mathematicians. In particular, several difficulties arise when extending the results and the methods from the scalar wave equation to vectorial problems modeling wave propagation in electromagnetism or elastodynamics.

A large part of research in mathematics, when applied to wave propagation problems, is oriented towards the following goals:

- The design of new numerical methods, increasingly accurate and efficient.
- The development of artificial transparent boundary conditions for handling unbounded propagation domains.
- The treatment of more and more complex configurations (non local models, non linear models, coupled systems, periodic media).
- The study of specific phenomena such as guided waves and resonances, which raise mathematical questions of spectral theory.
- The development of approximate models via asymptotic analysis with multiple scales (thin layers, boundary layers effects, small heterogeneities, homogenization, ...).
- The development and the analysis of algorithms for inverse problems (in particular for inverse scattering problems) and imaging techniques, using data from wave phenomena.

3.2. New schemes for time-domain simulations

Problems of wave propagation naturally arise as problems of evolution and it is necessary to have efficient methods for the calculation of their solution, directly in the time domain. The development and analysis of such methods has been in the past an important part of POEMS activity. Nowadays, there exists a large variety of higher order numerical methods that allow us to solve with good accuracy and in short computational time most classical wave propagation problems.

However, when one wishes to deal with real life applications, one has to tackle problems which are complex in many ways: they involve multi-physics, non standard (possibly nonlinear) constitutive laws, highly heterogeneous media with high contrasts of coefficients, complex geometries... In many cases, such problems escape to the direct application of the above mentioned methods and *ad hoc* dedicated methods have to be designed.

Such methods are most often of hybrid nature, which includes domain decomposition methods and subgridding, mixing of integral equations and PDEs, and artificial boundary conditions. In time domain, a particularly challenging issue is the time stability, in particular concerning the coupling of algorithms. To cope with this major difficulty, a key issue (and a kind of Grail for numerical analysts) is the development of energy preserving methods which is one of the specificity of the research developed at POEMS in this field.

3.3. Integral equations

Our activity in this field aims at developing accurate and fast methods for 3D acoustic and elastodynamic problems based on the discretization of boundary integral equations.

In traditional implementation, the dimensional advantage of Boundary Element Methods (BEM) with respect to domain discretization methods is offset by the fully-populated nature of the BEM matrix. Various approaches such as the Fast Multipole Method (FMM) or hierarchical matrices (H-matrices) have been proposed to overcome this drawback and derive fast BEMs. The specificity of our work consists in deriving such approaches not only for 3D acoustic wave propagation but also for 3D elastodynamics with applications in soil-structure interaction, seismology or seismic imaging.

Since the solution is computed through an iterative solver, a crucial point is then to control the number of iterations as the problem complexity increases, through the development of adapted preconditioners.

Besides, we also try to hybridize integral equations and high-frequency methods for scattering problems ,in order to tackle configurations with scatterers of different size-scales, compared to the wavelength.

Finally, we have studied the relationship between the Maxwell and eddy current models for three-dimensional configurations involving highly-conducting bounded bodies in air and sources placed remotely from those bodies

3.4. Domain decomposition methods

This is a come back to a topic in which POEMS contributed in the 1990's. It is motivated by our collaborations with the CEA-CESTA and the CEA-LIST, for the solution of large problems in time-harmonic electromagnetism and elastodynamics.

We combine in an original manner classical ideas of Domain Decomposition Methods with the specific formulations that we use for wave problems in unbounded domains, taking benefit of the available analytical representations of the solution (integral representation, modal expansion etc...).

One ANR project (NonLocalDD) supports this research.

3.5. Wave propagation in complex media

Our objective is first to develop efficient numerical approaches for the propagation of waves in heterogeneous media, taking into account their complex microstructure.

We aim on one hand to improve homogenized modeling of periodic media, by deriving enriched boundary conditions (or transmission conditions if the periodic structure is embedded in a homogeneous matrix) which take into account the boundary layer phenomena. On the other hand, we like to develop multi-scale numerical methods when the assumption of periodicity on the spatial distribution of the heterogeneities is relaxed, or even completely lost. The general idea consists in a coupling between a macroscopic solver, based on a coarse mesh, with some microscopic representation of the field. This latter can be obtained by a numerical microscopic solver or by an analytical asymptotic expansion. This leads to two very different approaches which may be relevant for very different applications.

Extraordinary phenomena regarding the propagation of electromagnetic or acoustic waves appear in materials which have non classical properties: materials with a complex periodic microstructure that behave as materials with negative physical parameters, metals with a negative dielectric permittivity at optical frequencies, magnetized plasmas endowed with a strongly anisotropic and sign-indefinite permittivity tensor. These non classical materials raise original questions from theoretical and numerical points of view.

3.6. Spectral theory and modal approaches

The study of waveguides is a longstanding and major topic of the team. Concerning the selfadjoint spectral theory for open waveguides, we turned recently to the very important case of periodic media. One objective is to design periodic structures with localized perturbations to create gaps in the spectrum, containing isolating eigenvalues.

Then, we would like to go further in proving the absence of localized modes in non uniform open waveguides. An original approach has been successfully applied to the scalar problem of a waveguides junctions or bent waveguides. The challenge now is to extend these ideas to vectorial problems (for applications to electromagnetism or elastodynamics) and to junctions of periodic waveguides.

Besides, we will continue our activity on modal methods for closed waveguides. In particular, we aim at extending the enriched modal method to take into account curvature and rough boundaries.

Finally, we are developing asymptotic models for networks of thin waveguides which arise in several applications (electric networks, simulation of lung, nanophotonics...).

3.7. Inverse problems

Building on the strong expertise of POEMS in the mathematical modeling of waves, most of our contributions aim at improving inverse scattering methodologies.

We acquired some expertise on the so called Linear Sampling Method, from both the theoretical and the practical points of view. Besides, we are working on topological derivative methods, which exploit small-defect asymptotics of misfit functionals and can thus be viewed as an alternative sampling approach, which take benefit of our expertise on asymptotic methods.

An originality of our activity is to consider inverse scattering in waveguides (the inverse scattering community generally considers only free-space configurations). This is motivated at the same time by specific issues concerning the ill-posedness of the identification process and by applications to non-destructive techniques, for waveguide configurations (cables, pipes, plates etc...). In particular, with the help of experimental data obtained at CEA-List, we proved the feasibility of the Linear Sampling Method to identify defects in the context of ultrasonic NDT.

Lastly, we continued our work on the so-called exterior approach for solving inverse obstacle problems, which associates quasi-reversibility and level set methods. We extended such approach to evolution problems, in particular the wave equation in the time domain for a finite time interval.

4. Application Domains

4.1. Acoustics

Two particular subjects have retained our attention recently.

1. Aeroacoustics, or more precisely, acoustic propagation in a moving compressible fluid, has been for our team a very challenging topic, which gave rise to a lot of open questions, from the modeling until the numerical approximation of existing models. Our works in this area are partially supported by Airbus. The final objective is to reduce the noise radiated by Airbus planes.
2. Musical acoustics constitute a particularly attractive application. We are concerned by the simulation of musical instruments whose objectives are both a better understanding of the behavior of existing instruments and an aid for the manufacturing of new instruments. We have successively considered the timpani, the guitar and the piano.

4.2. Electromagnetism

Applied mathematics for electromagnetism during the last ten years have mainly concerned stealth technology and electromagnetic compatibility. These areas are still motivating research in computational sciences (large scale computation) and mathematical modeling (derivation of simplified models for multiscale problems). Electromagnetic propagation in non classical media opens a wide and unexplored field of research in applied mathematics. This is the case of wave propagation in photonic crystals, metamaterials or magnetized plasmas. Finally, the simulation electromagnetic (possibly complex, even fractal) networks is motivated by destructive testing applications. These topics are developed in collaboration with CEA, DGA and ONERA.

4.3. Elastodynamics

Wave propagation in solids is with no doubt, among the three fundamental domains that are acoustics, electromagnetism and elastodynamics, the one that poses the most significant difficulties from mathematical and numerical points of view. A major application topic has emerged during the past years : the non destructive testing by ultra-sounds which is the main topic of our collaboration with CEA-LIST. On the other hand, we are developing efficient integral equation modelling for geophysical applications (soil-structure interaction for civil engineering, seismology).

5. Highlights of the Year

5.1. Highlights of the Year

- POEMS project-team reached the deadline of 12 years at the end of 2017. We have devoted a large part of our time during the first half-year to conceive and write a text of 20 pages which describes the new project that we submit to the management of Inria, in order to pursue our research on the modeling and simulation of wave phenomena. This project is currently discussed by several experts, in interaction with ourselves, before the final decision of creation of the new project-team.
- S. Chaillat co-organized with X. Claeys (Sorbonnes & EPI ALPINES) the symposium of the *International Association for Boundary Element Methods (IABEM)*, which took place in Paris in June 2018. There were about 140 attendees.
- A.-S. Bonnet-Ben Dhia co-organized a workshop entitled *“Advanced Theoretical and Numerical Methods for waves in structured Media”* in Paris in March 2018, in the framework of the GDR Ondes. There were about 90 attendees.
- P. Ciarlet is co-author of a book entitled *"Mathematical Foundations of Computational Electromagnetism"*, published in the serie *Applied Mathematical Sciences* by Springer.

6. New Software and Platforms

6.1. XLiFE++

KEYWORDS: Numerical simulations - Finite element modelling - Boundary element method
FUNCTIONAL DESCRIPTION: XLiFE++ is an FEM-BEM C++ code developed by POEMS laboratory and IRMAR laboratory, that can solve 1D/2D/3D, scalar/vector, transient/stationary/harmonic problems. Description: https://uma.ensta-paristech.fr/soft/XLiFE++/

- Contact: Eric Lunéville

6.2. COFFEE

KEYWORDS: Numerical simulations - Wave propagation - Boundary element method
FUNCTIONAL DESCRIPTION: COFFEE is an adapted fast BEM solver to model acoustic and elastic wave propagation (full implementation in Fortran 90). The 3-D acoustic or elastodynamic equations are solved with the boundary element method accelerated by the multi-level fast multipole method or a hierarchical-matrices based representation of the system matrix. The fundamental solutions for the infinite space are used in this implementation. A boundary element-boundary element coupling strategy is also implemented so multi-region problems (strata inside a valley for example) can be solved. In order to accelerate the convergence of the iterative solver, various analytic or algebraic preconditioners are available. Finally, an anisotropic mesh adaptation strategy is used to further reduce the computational times.

- Contact: Stéphanie Chaillat

7. New Results

7.1. New schemes for time-domain simulations

7.1.1. Solving the Isotropic Linear Elastodynamics Equations Using Potentials

Participant: Patrick Joly.

This work is done in collaboration with Sébastien Impériale (EPI M3DISIM) and Jorge Albella and Jeronimo Rodríguez from the University of Santiago de Compostela.

We pursue our research on the numerical solution of 2D elastodynamic equations in piecewise homogeneous media using the decomposition of the displacement fields into the sum of the gradient and the rotational (respectively) of two scalar potentials potentials. This allows us to obtain an automatic decomposition of the wave field into the sum of pressure and shear waves (respectively). The approach is expected to be efficient when the velocity of shear waves is much smaller than the velocity of pressure waves, since one can adapt the discretization to each type of waves. This appears as a challenge for finite element methods , the most delicate issue being the treatment of boundary and transmission conditions, where the two potentials are coupled..

A stable (mixed) variational formulation of the evolution problem based on a clever choice of Lagrange multipliers has been proposed as well as various finite element approximations which have been successfully implemented. The analysis of the continuous problem has been published in a long paper in the journal of Scientific computing.The numerical analysis of the discretized problem is in progress.

7.1.2. Time domain Half-Space Matching method

Participants: Sonia Fliss, Hajer Methenni.

This work is done in the framework of the PhD of Hajer Methenni (funded by CEA-LIST) and in collaboration with Sebastien Imperiale (EPI M3DISIM) and Alexandre Imperiale (CEA-LIST).

The objective of this work is to propose a numerical method to solve the elastodynamics equations in a locally perturbed unbounded anisotropic media. Let us mention that all the classical methods to restrict the computation around the perturbations are unstable in anisotropic elastic media (PMLs for instance) or really costly (Integral equations). The idea is to extend the method already developed for the corresponding time harmonic problem, called the Halfspace Matching Method. We have considered, for now, the 2D scalar wave equation but the method is constructed in order to be applied to the elastodynamic problem. The method consists in coupling several representations of the solution in half-planes surrounding the defect with a FE representation in a bounded domain including the defect. In order to ensure the stability of the method, we first semi-discretize in time the equations and apply the method to the semi-discrete problem. Thus, for each time step, by ensuring that all the representations of the solution match, in particular in the intersection of the half-planes, we end up, at each time step, with a system of equations which couples, via integral operators, the solution at this time step in the bounded domain and its traces on the edge of the half-planes, the right hand side being a convolution operator involving the solution at the previous time steps. The method has been implemented and validated with Xlife++.

We are now looking to make the method more efficient by implementing methods of acceleration. Finally, we will also seek to develop another version of the method based on the Convolution quadrature.

7.1.3. Time domain modelling for wave propagation in fractal trees

Participants: Patrick Joly, Maryna Kachanovska.

In order to simulate wave propagation in fractal trees (see section 7.4.3), which have infinite structure, it is necessary to be able to truncate the computations to a finite subtree. This was done using Dirichlet-to-Neumann (DtN) operators in our previous work in collaboration with A. Semin (TU Darmstadt). In this case a DtN operator is a convolution operator, whose kernel is not known in a closed form. Based on the results of this previous work, in 2017 we had proposed two methods for approximating these convolution operators:

- constructing an exact DtN operator for a semi-discretized system (in the spirit of convolution quadrature methods).
- truncating meromorphic expansion for the symbol (Fourier transform of the convolution kernel) of the DtN operator, which allows to approximate the DtN operator by local operators.

This year we have performed a complete convergence and stability analysis of these methods, based on the energy techniques.

In particular, for the convolution quadrature methods, we were able to obtain all the estimates using time-domain analysis, by avoiding passage to the Laplace domain.

As for the method based on the meromorphic expansion of the symbol of the DtN operator, we have shown that the error induced by truncating the expansion to L terms can be controlled by a remainder of a series, which, in particular, depends on the eigenvalues of the weighted Laplacian on the fractal trees. To obtain an explicit dependence of the error on L, we have computed Weyl bounds for the eigenvalues, based on a refinement of the ideas of [Kigami, Lapidus, Comm. Math. Phys. 158 (1993)].

Additionally, we have addressed some computational aspects of the two methods, in particular, efficient evaluation of the symbol of the DtN operator (we have an algorithm that allows to evaluate it at the frequency ω in $O(\log^k |\omega|)$ time), as well as a method for efficient computation of the poles of the symbol (based on Möbius transform and polynomial interpolation).

7.2. Integral equations and boundary element methods (BEMs)

7.2.1. Accelerated and adapted BEMs for wave propagation

Participants: Faisal Amlani, Stéphanie Chaillat.

This work is done in collaboration with Adrien Loseille (EPI Gamma3).

We extend to high-order curved elements a recently introduced metric-based anisotropic mesh adaptation strategy for accelerated boundary element methods (e.g. Fast Multipole(FM-) BEM) applied to exterior boundary value problems. This method derives from an adaptation framework for volumetric finite element methods and is based on an iterative procedure that completely remeshes at each refinement step and that leads to a strategy that is independent of discretization technique (e.g., collocation or Galerkin) and integral representation (e.g., single- or double-layer). In effect, it results in a truly anisotropic adaptation that alters the size, shape and orientation of each element according to an optimal metric based on a numerically recovered Hessian of the boundary solution. The algorithm is principally characterized by its ability to recover optimal convergence rates for both flat and curved discretizations (e.g. P_0-, P_1- or P_2-elements) of a geometry containing singularities such as corners and edges. This is especially powerful for realistic geometries that include engineering detail (whose solutions often entail severe singular behavior).

Additionally, we address — by way of introducing hierarchical ($\mathcal{H}$-) matrix preconditioning applied to fast multipole methods via a Flexible GMRES (FGMRES) routine — the computational difficulties that arise when resolving highly anisotropic (and hence highly ill-conditioned) linear systems. The new technique, which uses a very coarse $\mathcal{H}$-matrix system (constructed rapidly via high-performance parallelization) to precondition the full Fast Multipole Method system, drastically reduces the overall computation time as well as the iterative solve time, further improving the tractability of addressing even larger and more complex geometries by FM-BEM.

7.2.2. *Preconditioned $\mathcal{H}$-matrix based BEMs for wave propagation*

Participants: Stéphanie Chaillat, Patrick Ciarlet, Félix Kpadonou.

We are interested with fast boundary element methods (BEMs) for the solution of acoustic and elastodynamic problems.

The discretisation of the boundary integral equations, using BEM, yields to a linear system, with a fully-populated matrix. Standard methods to solve this system are prohibitive in terms of memory requirements and solution time. Thus one is rapidly limited in terms of complexity of problems that can be solved. The $\mathcal{H}$-matrix based BEMs is commonly used to address these limitations. It is a purely algebraic approach.

The starting point is that the BEM matrix can be partitioned into some blocks which can either be of low or full rank. Memory can be saved by using low-rank revealing technique such as the Adaptive Cross Approximation. We have already study the efficiency of this approach for wave propagation problems. The purpose being the applications to large scale problems, we are now interested in an efficient implementation of the solver in a high performance computing setting. Thus, a bottleneck, with an hierarchical matrix data-sparse representation, is the management of the memory and its (prior) estimation for array allocations.

The first part of our work has been devoted to the proposition of an a priori estimation of the ranks of the blocks in the hierarchical matrix. Afterwards, we have implemented a parallel construction of the $\mathcal{H}$-matrix representation and H-matrix vector product (basic operation in any iterative solver), using a multi-threading OpenMP parallelization. The solution is then computed through the GMRES iterative solver. A crucial point is then the solution time of that solver and the number of iterations as the problem complexity increases. We have developed a two-level, nested outer-inner, iterative solver strategy. The inner solver preconditioned the outer. The preconditioner is a coarse data-sparse representation of the BEM system matrix.

7.2.3. *Coupling integral equations and high-frequency methods*

Participants: Marc Bonnet, Marc Lenoir, Eric Lunéville, Laure Pesudo.

This theme concerns wave propagation phenomena which involve two different space scales, namely, on the one hand, a medium scale associated with lengths of the same order of magnitude as the wavelength (medium-frequency regime) and on the other hand, a long scale related to lengths which are large compared to the wavelength (high-frequency regime). Integral equation methods are known to be well suited for the former, whereas high-frequency methods such as geometric optics are generally used for the latter. Because of the presence of both scales, both kinds of simulation methods are simultaneously needed but these techniques do not lend themselves easily to coupling.

The scattering of an acoustic wave by two sound-hard obstacles: a large obstacle subject to high-frequency regime relatively to the wavelength and a small one subject to medium-frequency regime has been investigated by Marc Lenoir, Eric Lunéville and Laure Pesudo. The technique proposed in this case consists in an iterative method which allows to decouple the two obstacles and to use Geometric Optics or Physical Optics for the large obstacle and Boundary Element Method for the small obstacle. This approach has been validated on various situations using the XLife++ library developed in the lab. When the obstacles are not sticked, even if they are very close, the iterative method coupling BEM and some high-frequency methods (ray approximation or Kirchoff approximation) works very well. When the obstacle are sticked, the "natural" iterative method is no longer convergent. We are currently looking for some improved methods to deal with these cases that have a practical interest.

7.2.4. The eddy current model as a low-frequency, high-conductivity asymptotic form of the Maxwell transmission problem

Participant: Marc Bonnet.

In this work, done in collaboration with Edouard Demaldent (CEA LIST), we study the relationship between the Maxwell and eddy current (EC) models for three-dimensional configurations involving highly-conducting bounded bodies in air and sources placed remotely from those bodies. Such configurations typically occur in the numerical simulation of eddy current non destructive testing (ECT). The underlying Maxwell transmission problem is formulated using boundary integral formulations of PMCHWT type. In this context, we derive and rigorously justify an asymptotic expansion of the Maxwell integral problem with respect to the non-dimensional parameter $\gamma := \sqrt{\omega\varepsilon_0/\sigma}$. The EC integral problem is shown to constitute the limiting form of the Maxwell integral problem as $\gamma \to 0$, i.e. as its low-frequency and high-conductivity limit. Estimates in γ are obtained for the solution remainders (in terms of the surface currents, which are the primary unknowns of the PMCHWT problem, and the electromagnetic fields) and the impedance variation measured at the extremities of the excitating coil. In particular, the leading and remainder orders in γ of the surface currents are found to depend on the current component (electric or magnetic, charge-free or not). Three-dimensional illustrative numerical simulations corroborate these theoretical findings.

7.2.5. Modelling the fluid-structure coupling caused by a far-field underwater explosion

Participants: Marc Bonnet, Stéphanie Chaillat, Damien Mavaleix-Marchessoux.

This work, funded by Naval Group and a CIFRE PhD grant, addresses the computational modelling of the mechanical effect on ships of remote underwater explosions. We aim at a comprehensive modelling approach that accounts for the effect of the initial (fast) wave impinging the ship as well as that of later, slower, water motions. Both fluid motion regimes are treated by boundary element methods (respectively for the wave and potential flow models), while the structure is modelled using finite elements. To cater for large and geometrically complex structures, the BEM-FEM interface requires large numbers of DOFs, which entails the use of a fast BEM solver. Accordingly, the wave-like fluid motions are to be computed by means of the convolution quadrature method (CQM) implemented in the in-house fast BEM code COFFEE. This work is in progress (the thesis having started in Dec. 2017). Work accomplished so far has mainly consisted in (a) thoroughly examinating the physical modelling issues, (b) formulating the mathematical and computational model that takes relevant physical features into account, and (c) implementing and assessing the CQM under conditions similar to those of the aimed application.

7.3. Domain decomposition methods

7.3.1. Transparent boundary conditions with overlap in unbounded anisotropic media

Participants: Anne-Sophie Bonnet Ben-Dhia, Sonia Fliss, Yohanes Tjandrawidjaja.

This work is done in the framework of the PhD of Yohanes Tjandrawidjaja (funded by CEA-LIST), in collaboration with Vahan Baronian (CEA). This follows the PhD of Antoine Tonnoir (now Assistant Professor at Insa of Rouen) who developed a new approach, the Half-Space Matching Method, to solve scattering problems in 2D unbounded anisotropic media. The objective is to extend the method to a 3D plate of finite width.

In 2D, our approach consists in coupling several plane-waves representations of the solution in half-spaces surrounding the defect with a FE computation of the solution around the defect. The difficulty is to ensure that all these representations match, in particular in the infinite intersections of the half-spaces. It leads to a formulation which couples, via integral operators, the solution in a bounded domain including the defect and some traces of the solution on the edges of the half-planes. We have proven that, in presence of dissipation, this system is a Fredholm equation of the second kind, in an L2 functional framework. The truncation of the Fourier integrals and the finite element approximation of the corresponding numerical method have been also analyzed.

The method has been extended to the 3D case, for an application to non-destructive testing. The objective is to simulate the interaction of Lamb waves with a defect in an anisotropic elastic plate. The additional complexity compared to the 2D case lies in the representations which are obtained semi-analytically by decomposition on Lamb modes. In addition, the system of equations couples the FE representation in the bounded perturbed domain with not only the displacement, but also the normal stress of the solution on the infinite bands limiting the half-plates. A first numerical result has been obtained in the isotropic case.

The perspectives now concern the efficiency of the method (which could be improved by replacing the direct inversion by a preconditioned iterative inversion with an efficient product matrix-vector), the analysis of the method in the case without dissipation and the analysis of the method in the elastic case.

7.3.2. Coupling BEMs in overlapping domains when a global Green's function is not available

Participants: Anne-Sophie Bonnet Ben-Dhia, Stéphanie Chaillat, Sonia Fliss, Yohanes Tjandrawidjaja.

We consider in this work problems for which the Green's function is not available, so that classical Boundary Integral equation methods are not applicable. Let us mention for instance the junction of two different stratified media (tapered optical fibers in integrated optics or junction of two topographic elastic surfaces in geophysics).

To this end, we propose a generalization of the Half-Space Matching method (see section 7.3.1).

In this work, by replacing the Fourier representations by integral representations, we are able to replace the half-spaces by more general unbounded overlapping sub-domains. We choose the sub-domains in such a way that an explicit Green's function is available for each subdomain. For instance, for the configuration described above (figure 1a), it suffices to introduce two infinite sub domains, each of them containing only one stratification (figures 1c and 1d) and a bounded domain containing the junction (figure 1b). The formulation couples the solution in the bounded domain with the single and double layer potentials on each boundary of the sub-domains. The approximation relies on a FE discretisation of the volume unknown and a truncation and a discretization of the boundary/surface unknowns.

A study concerning the choice of the discretisation parameters and the shape of the infinite lines have to be done. The theoretical analysis of the method raises challenging open questions: for instance, a first uniqueness result has been derived, which requires the definition of a variational formulation on a Rie. Finally, we want to apply the method to the scattering by a step, i.e. the junction of two semi infinite-planes joined together by a step.

7.3.3. Domain decomposition method for acoustics with uniform exponential rate of convergence using non-local impedance operators

Participants: Patrick Joly, Francis Collino, Émile Parolin.

This work is done in the framework of the PhD of Émile Parolin (funded by ANR NonlocalDD), in collaboration with X. Clayes (EPI Alpines & LJLL).

We continued the work on non-overlapping domain decomposition methods with non-local transmission conditions for time-harmonic wave propagation. The analysis of such methods is conducted by writing them as a relaxed Jacobi algorithm. In the absence of junctions points, the continuous algorithm converges exponentially fast under suitable assumptions on the impedance operators. These assumptions cannot be satisfied using local operators and rely in practice on singular integral operators. The progress achieved is as follows.

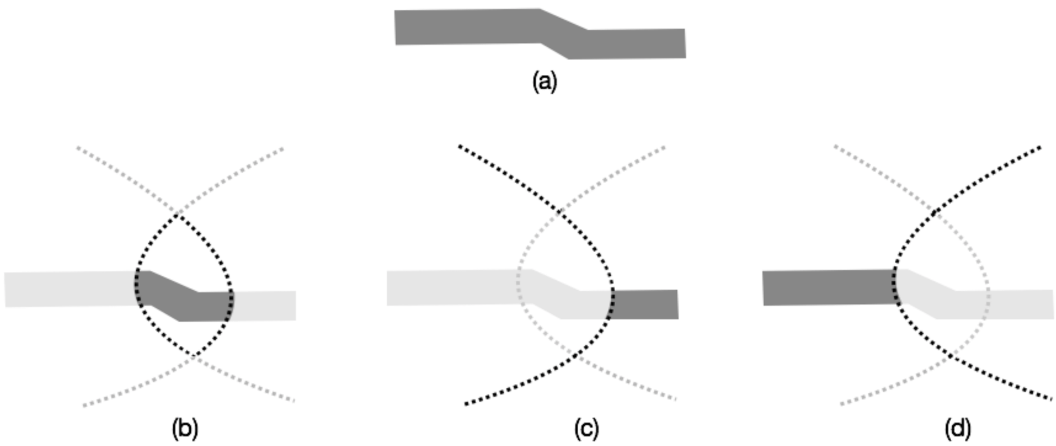


Figure 1. Coupling BEMs in overlapping domains

- In the context of acoustic wave propagation, we established a new result on the robustness of the algorithm with respect to the mesh size. We have proven that for Lagrange finite element approximations the exponential rate of convergence of the algorithm is independent of the discretization parameter, hence does not deteriorate when the mesh is refined. The proof relies on the Scott-Zhang interpolator and led to the submission.
- We have been working on the extension to 3D time harmonic Maxwell's equations. The main difficulty is to design well adapted operators taking into account the specificity of the corresponding trace spaces. An adequate operator must behave like a pseudo-differential operator with opposite order on the 'curl part' and 'grad part' of a tangential field (this is related to the Helmholtz decomposition of tangential fields). Guided by potential theory for elliptic operators, we proposed two classes of suitable operators. The first one is based on Bessel potentials (fractional powers of the shifted Laplacian) and the second one relies on Riesz potentials. We have shown that the proposed operators satisfy the desired properties in the case of a sphere using modal analysis techniques. We have also been working on the design of the finite element approximation of these operators.

7.3.4. Domain decomposition method with cross-point treatment for high-frequency acoustic scattering

Participant: Axel Modave.

This work is done in collaboration with X. Antoine (IECL & EPI SPHINX) and C. Geuzaine (Université de Liège).

Solving high-frequency time-harmonic scattering problems using FE techniques is challenging, as such problems lead to very large, complex and indefinite linear systems. Optimized Schwarz domain decomposition methods (DDMs) are currently a very promising approach, where subproblems of smaller sizes are solved in parallel using direct solvers, and are combined in an iterative procedure. It is well-known that the convergence rate of these methods strongly depends on the transmission condition enforced on the interfaces between the subdomains.

Local transmission conditions based on high-order absorbing boundary conditions (HABCs) have proved well suited. They represent a good compromise between basic impedance conditions (which lead to suboptimal convergence) and the exact Dirichlet-to-Neumann (DtN) map related to the complementary of the subdomain (which is expensive to compute). However, a direct application of this approach for domain decomposition configurations with cross-points, where more than two subdomains meet, does not provide satisfactory results.

We work on improved DDMs that efficiently addresses configurations with cross-points. Noting that these points actually are corners for the subdomains, our strategy consists in incorporating a corner treatment developed for HABCs (see section 7.7.1) into the DDM procedure. We propose a cross-point treatment for HABC-based DDMs in settings with cross-points and right angles. The method is implemented and successfully tested for two-dimensional examples. The analysis of this method is currently in progress. Extensions to more complicated settings (e.g. 3D, with non-right angles, other physical waves) will be investigated in the future.

7.4. Wave propagation in complex media

7.4.1. Enriched Homogenization in presence of boundaries or interfaces

Participants: Clement Beneteau, Sonia Fliss.

This work is done in the framework of the PhD of Clement Beneteau and is done in collaboration with X. Claeys (Sorbonne & EPI Alpines).

This work is motivated by the fact that classical homogenization theory poorly takes into account interfaces or boundaries. It is particularly unfortunate when one is interested in phenomena arising at the interfaces or the boundaries of the periodic media (the propagation of plasmonic waves at the surface of metamaterials for instance). To overcome this limitation, we have constructed an effective model which is enriched near the interfaces and/or the boundaries. For now, we have treated and analysed the case of simple geometries: for instance a half-plane with Dirichlet or Neumann boundary conditions or a plane interface between two periodic half spaces. We have derived a high order approximate model which consists in replacing the periodic media by an effective one but the boundary/transmission conditions are not classical. The obtained conditions involve Laplace- Beltrami operators at the interface and requires to solve cell problems in periodicity cell (as in classical homogenization) and in infinite strips (to take into account the phenomena near the boundary/interface). We establish well posedness for the approximate model and error estimates which justify that this new model is more accurate. From a numerical point of view, the only difficulty comes from the problems set in infinite strips. The method has been implemented using Xlife++.

This approach has been extended to the long time homogenisation of the wave equation. It is well known that the classical effective homogenized wave equation does not capture the long time dispersive effects of the waves in the periodic media. Since the works of Santosa and Symes in the 90's, several effective equations (involving differential operators of order at least 4) that capture these dispersive effects have been proposed, but only in infinite media. In presence of boundaries or interfaces, the question of boundary/transmission conditions for these effective equations was never treated. We have first results in that direction.

7.4.2. Transmission conditions between homogeneous medium and periodic cavities

Participant: Jean-François Mercier.

In collaboration with A. Maurel (Langevin Institute), J. J. Marigo (LMS) and K. Pham (Imsia).

We have developed a model for resonant arrays of Helmholtz cavities, thanks to a two scale asymptotic analysis. The model combines volumic homogenization to replace the cavity region by a homogeneous anisotropic slab and interface homogenization to replace the region of the necks by transmission conditions. The coefficients entering in the effective wave equation are simply related to the fraction of air in the periodic cell of the array. Those involved in the jump conditions encapsulate the effects of the neck geometry.

In parallel, this effective model has been exploited to study the resonance of the Helmholtz resonators with a focus on the influence of the neck shape. The homogenization makes a parameter B to appear which determines unambiguously the resonance frequency of any neck. As expected, this parameter depends on the length and on the minimum opening of the neck, and it is shown to depend also on the surface of air inside the neck. Once these three geometrical parameters are known, B has an additional but weak dependence on the neck shape, with explicit bounds.

7.4.3. *Mathematical analysis of wave propagation in fractal trees*

Participants: Patrick Joly, Maryna Kachanovska.

We have continued our work (in collaboration with A. Semin (TU Darmstadt)) on wave propagation in fractal trees which model human lungs. One of the major results of this year is a complete analysis of such models. In particular, provided Sobolev spaces H^1_μ, L^2_μ (which generalize weighted Sobolev spaces on an interval to the case of fractal trees) we clarified the following questions for a range of parameters of the trees not covered by the previous theory: existence of traces of H^1_μ-functions on fractal trees; approximation of H^1_μ-functions by compactly supported functions; compact embedding of H^1_μ into L^2_μ.

7.4.4. *Hyperbolic Metamaterials in Frequency Domain: Free Space*

Participants: Patrick Ciarlet, Maryna Kachanovska.

In this project we consider the wave propagation in 2D hyperbolic metamaterials [Poddubny et al., Nature Photonics, 2013], which are modelled by Maxwell equations with a diagonal frequency-dependent tensor of dielectric permittivity ε and scalar frequency-independent magnetic permeability. In the time domain, the corresponding models are well-posed and stable. Surprisingly, in some regimes in the frequency domain, when the signs of the diagonal entries of ε do not coincide, the problem becomes hyperbolic (and hence the name). The main goal of this project is to justify the well-posedness of such models in the frequency domain, first of all starting with the case of the free space. We have obtained partial results in this direction: radiation condition, which ensures the well-posedness of the problem, mapping properties of the resolvent (with refined estimates on the propagation of singularities in these models). We are currently working on the limiting absorption and limiting amplitude principles.

7.5. Spectral theory and modal approaches for waveguides

7.5.1. *Scattering solutions in an unbounded strip governed by a plate model*

Participants: Laurent Bourgeois, Sonia Fliss.

Together with Lucas Chesnel (EPI DEFI), we have initiated a new work on a particular waveguide which consists of a thin strip governed by a Kirchhoff-Love bilaplacian model. The aim is to build some radiation conditions and prove well-posedness of scattering problems for that simple model and for two kinds of boundary conditions: the strip is either simply supported or clamped. In the first case, we have shown that using a Dirichlet-to-Neumann operator enables us to prove fredholmness. Such approach is not possible in the second case, for which a completely different angle of attack is chosen: a Kondratiev approach involving weighted Sobolev spaces and detached asymptotics.

7.5.2. *Modal analysis of electromagnetic dispersive media*

Participants: Christophe Hazard, Sandrine Paolantoni.

We investigate the spectral effects of an interface between vacuum and a negative material (NM), that is, a dispersive material whose electric permittivity and/or magnetic permeability become negative in some frequency range. Our first work in this context concerns an elementary situation, namely, a two-dimensional scalar model (derived from the complete Maxwell's equations) which involves the simplest existing model of NM, referred to as the non-dissipative Drude model (for which negativity occurs at low frequencies). By considering a polygonal cavity, we have shown that the presence of the Drude material gives rise to various components of an essential spectrum corresponding to various unusual resonance phenomena: first, a low frequency bulk resonance (accumulation at the zero frequency of positive eigenvalues whose associated eigenvectors are confined in the Drude material); then, a surface resonance for one particular critical frequency (at which the so-called surface plasmons occurs, that is, localized highly oscillating vibrations at the interface between the Drude material and the vacuum); finally, corner resonances in a critical frequency interval (here, localized highly oscillating vibrations occur near any corner of the interface, interpreted as a "black hole" phenomenon). An article which presents these results has been submitted. Most recent works were devoted to the numerical simulation of these resonance phenomena in the context of the code XLiFE++ developped in the lab.

7.5.3. Formulation of invisibility in waveguides as an eigenvalue problem

Participant: Anne-Sophie Bonnet-Ben Dhia.

This work is done in collaboration with Lucas Chesnel from EPI DEFI and Vincent Pagneux from Laboratoire d'Acoustique de l'Université du Maine.

We consider an infinite acoustic waveguide (with a bounded cross-section) which is locally perturbed. At some exceptional frequencies and for particular incident waves, it may occur that all the energy of the incident wave is transmitted, the only effect in reflection being a superposition of evanescent modes in the vicinity of the perturbation. We have proposed an approach for which these reflection-less frequencies appear directly as eigenvalues of a new problem. This problem is very similar to the formulation of the scattering problem using Perfectly Matched Layers, except a slight modification in the PML. Precisely, we use two conjugated dilation parameters, α in the outlet and $\overline{\alpha}$ in the inlet, in order to select outgoing waves in the outlet and ingoing waves in the inlet. In fact, we show that the real eigenfrequencies that are obtained correspond either to trapped modes or to reflection-less modes. In addition to this real spectrum, we find intrinsic complex frequencies, which also contain information about the quality of the transmission through the waveguide. Mathematically, the non-selfadjoint eigenvalue problem with conjugated PMLs has strange properties: the discreteness of the point spectrum is not stable by compact perturbations and pathological examples can be exhibited.

7.6. Inverse problems

7.6.1. Linear Sampling Method with realistic data in waveguides

Participants: Laurent Bourgeois, Arnaud Recoquillay.

Our activities in the field of inverse scattering in waveguides with the help of sampling methods has now a quite long history. Very recently, we have focused on elastodynamics and realistic data, that is surface data in the time domain. This has been the subject of the PhD of Arnaud Recoquillay. It was motivated by Non Destructive Testing activities for tubular structures and was the object of a partnership with CEA List (Vahan Baronian).

Our strategy consists in transforming the time domain problem into a multi-frequency problem by the Fourier transform. This allows us to take full advantage of the established efficiency of modal frequency-domain sampling methods. In particular, we have shown how to optimize the number of sources/receivers and the distance between them in order to obtain the best possible imaging results.

Our main achievement is an experimental validation of such approach in the presence of real data: the measurements were carried at CEA on steel plates with the help of piezoelectric sensors. The identification results are encouraging and pave the way of a future integration of sampling methods in real NDT activities.

7.6.2. The "exterior approach" to solve inverse obstacle problems

Participants: Laurent Bourgeois, Arnaud Recoquillay, Dmitry Ponomarev.

This work is done in collaboration with Jérémi Dardé (IMT Toulouse).

We consider some inverse obstacle problems in acoustics by using a single incident wave, either in the frequency or in the time domain. When so few data are available, a Linear Sampling type method cannot be applied. In order to solve those kinds of problem, we propose an "exterior approach", coupling a mixed formulation of quasi-reversibility and a simple level set method. In such iterative approach, for a given defect D, we update the solution u with the help of a mixed formulation of quasi-reversibility while for a given solution u, we update the defect D with the help of a level set method based on a Poisson problem. We have studied two cases. The first case concerns the waveguide geometry in the frequency domain. The second case concerns a bounded spatial set in the time domain when data are given in a finite time interval. This last case is challenging because it raises the (open) question of the minimal final time which is required to ensure uniqueness of the obstacle from the lateral Cauchy data.

7.6.3. Inverse acoustic scattering using high-order small-inclusion expansion of misfit function

Participant: Marc Bonnet.

This work concerns an extension of the topological derivative concept for 3D inverse acoustic scattering problems involving the identification of penetrable obstacles, whereby the featured data-misfit cost function J is expanded in powers of the characteristic radius a of a single small inhomogeneity. The $O(a^6)$ approximation of J is derived and justified for a single obstacle of given location, shape and material properties embedded in a 3D acoustic medium of arbitrary shape, and the generalization to multiple small obstacles is outlined. Simpler and more explicit expressions are obtained when the scatterer is centrally-symmetric or spherical. An approximate and computationally light global search procedure, where the location and size of the unknown object are estimated by minimizing the $O(a^6)$ approximation over a search grid, is proposed and demonstrated on numerical experiments, where the identification from known acoustic pressure on the surface of a penetrable scatterer embedded in a acoustic semi-infinite medium, and whose shape may differ from that of the trial obstacle assumed in the expansion of J, is considered. measurements configuration situated far enough from the probing region.

7.6.4. Microstructural topological sensitivities of the second-order macroscopic model for waves in periodic media

Participant: Marc Bonnet.

This work is done in collaboration with Bojan Guzina (University of Minnesota, USA) and Rémi Cornaggia (IRMAR, Rennes).

We consider scalar waves in periodic media through the lens of a second-order effective i.e. macroscopic description, and we aim to compute the sensitivities of the relevant effective parameters due to topological perturbations of a microscopic unit cell. Specifically, our analysis focuses on the tensorial coefficients in the governing mean-field equation – including both the leading order (i.e. quasi-static) terms, and their second-order counterparts. The results demonstrate that the sought sensitivities are computable in terms of (i) three unit-cell solutions used to formulate the unperturbed macroscopic model; (ii) two adoint-field solutions driven by the mass density variation inside the unperturbed unit cell; and (iii) the usual polarization tensor, appearing in the related studies of non-periodic media, that synthesizes the geometric and constitutive features of a point-like perturbation. The proposed developments may be useful toward (a) the design of periodic media to manipulate macroscopic waves via the microstructure-generated effects of dispersion and anisotropy, and (b) sub-wavelength sensing of periodic defects or perturbations.

7.6.5. Analysis of topological derivative as a tool for qualitative identification

Participant: Marc Bonnet.

This work is a collaboration with Fioralba Cakoni (Rutgers University, USA).

The concept of topological derivative has proved effective as a qualitative inversion tool for a wave-based identification of finite-sized objects. Although for the most part, this approach remains based on a heuristic interpretation of the topological derivative, a first attempt toward its mathematical justification was done in Bellis et al. (Inverse Problems 29:075012, 2013) for the case of isotropic media with far field data and inhomogeneous refraction index. Our paper extends the analysis there to the case of anisotropic scatterers and background with near field data. Topological derivative-based imaging functional is analyzed using a suitable factorization of the near fields, which became achievable thanks to a new volume integral formulation recently obtained in Bonnet (J. Integral Equ. Appl. 29:271-295, 2017). Our results include justification of sign heuristics for the topological derivative in the isotropic case with jump in the main operator and for some cases of anisotropic media, as well as verifying its decaying property in the isotropic case with near field spherical measurements configuration situated far enough from the probing region.

7.6.6. Elasticity imaging by error in constitutive equation functionals

Participant: Marc Bonnet.

This work is done in collaboration with Wilkins Aquino (Duke University, USA).

We formulate the identification of heterogeneous linear elastic moduli in the context of time-harmonic elastodynamics as the minimization of the modified error in constitutive equation (MECE) functional. Our main goal is to develop theoretical foundations, in a continuous setting, allowing to explain and justify some known beneficial properties of this treatment. A specific feature of MECE formulations is that forward and adjoint solutions are governed by a fully coupled system, whose mathematical properties play a fundamental role in the qualitative and computational aspects of MECE minimization. We prove that this system has a unique and stable solution at any frequency, provided data is abundant enough (in a sense made precise), even though the relevant forward problem is not *a priori* clearly defined. This result has practical implications such as applicability of MECE to partial interior data (with important practical applications including ultrasound elastography), convergence of finite element discretizations and differentiability of the reduced MECE functional. In addition, we establish that usual least squares and pure ECE formulations are limiting cases of MECE formulations for small and large values of the weight of the data misfit component of the functional, respectively. For the latter case, we furthermore show that the reduced MECE Hessian is asymptotically positive for any parameter perturbation supported on the measurement region, thereby corroborating existing computational evidence on convexity improvement brought by MECE functionals. Finally, numerical studies including parameter reconstruction examples using interior data support our findings.

7.6.7. A continuation method for building large invisible obstacles in waveguides

Participants: Antoine Bera, Anne-Sophie Bonnet-Ben Dhia.

This work is done on collaboration with Lucas Chesnel (EPI DEFI).

We are interested in building invisible obstacles in waveguides, at a given frequency. The invisibility is characterized by the nullity of the scattering coefficients associated to propagating modes. In previous papers, a method has been proposed to prove the existence of invisible obstacles and to build them. But its main drawback was its limitation to small obstacles. In order to get larger invisible obstacles, we have developed a new approach which combines the previous idea with a continuation method: we are building a sequence of invisible obstacles, each of them being a small perturbation of the previous one. This algorithm is based, at each step, on the ontoness of an application and on the fixed-point theorem. We have implemented the method in the finite element library XLiFE++, in the case of penetrable obstacles of a two-dimensional acoustic waveguide, in multi-modal regime. A remarkable result is that the ontoness condition can be ensured in many cases, so that the algorithm can be iterated as long as required. Another interesting feature of our approach is that it allows to prescribe some properties of the obstacle (shape of the obstacle, piecewise constant index, ...), but a drawback is that the algorithm can produce non-realistic negative indices. This is a question that we are currently working on. Finally, let us emphasize that the formalism of the method is very general and flexible. In particular, it can be directly extended to 3D waveguides, or to the scattering in free space.

7.7. Acoustics and aeroacoustics

7.7.1. High-order absorbing boundary conditions with corner treatment for high-frequency acoustic scattering

Participant: Axel Modave.

This work is done in collaboration with C. Geuzaine (University of Liège) and X. Antoine (IECL & EPI SPHINX)

We address the design and validation of accurate local absorbing boundary conditions set on convex polygonal computational domains for the finite element solution of high-frequency acoustic scattering problems. While high-order absorbing boundary conditions (HABCs) are accurate for smooth fictitious boundaries, the precision of the solution drops in the presence of corners if no specific treatment is applied. We analyze two strategies to preserve the accuracy of Padé-type HABCs at corners: first by using compatibility relations (derived for right angle corners) and second by regularizing the boundary at the corner. We show that the former strategy is well-adapted to right corners and efficient for nearly-right corners, while the later is better for very obtuse corners. Numerical results are proposed to analyze and compare the approaches for two-and three-dimensional problems.

7.7.2. Time-harmonic acoustic scattering in a vortical flow

Participants: Antoine Bensalah, Patrick Joly, Jean-François Mercier.

We study the time-harmonic acoustic radiation in a fluid in flow. To go beyond the convected Helmholtz equation, only adapted to potential flows, we use Goldstein's equations, coupling exactly the acoustic waves to the hydrodynamic field. We have studied the hydrodynamic part of Goldstein equations, corresponding to a generalized time-harmonic transport equation and we have investigated its well-posedness. The result has been established under the assumption of a domain-filling flow, which in 2D is simply equivalent to a flow that does not vanish. The approach relies on the method of characteristics, which leads to the resolution of the transport equation along the streamlines and on general results of functional analysis. The theoretical results have been illustrated with numerical results obtained with a SUPG Finite Element scheme.

In complement we have developed a new model for Goldstein's equations in which the description of the hydrodynamic phenomena is simplified. The model, initially developed for a carrier flow of low Mach number M, is proved theoretically to remain accurate for moderate Mach numbers, associated to a low error bounded by M^2. Numerical experiments confirm the M^2 law and the good quality of the model for flows of non-small Mach numbers.

7.8. Numerical analysis for PDEs

7.8.1. A family of Crouzeix-Raviart Finite Elements in 3D

Participant: Patrick Ciarlet.

This work is done in collaboration with C. Dunkl (University of Virginia) and S. Sauter (Universität Zürich).

We develop a family of non-conforming "Crouzeix–Raviart" type finite elements in three dimensions. They consist of local polynomials of maximal degree p on simplicial finite element meshes while certain jump conditions are imposed across adjacent simplices. We will prove optimal a priori estimates for these finite elements. The characterization of this space via jump conditions is implicit and the derivation of a local basis requires some deeper theoretical tools from orthogonal polynomials on triangles and their representation. We will derive these tools for this purpose. These results allow us to give explicit representations of the local basis functions. Finally, we will analyze the linear independence of these sets of functions and discuss the question whether they span the whole non-conforming space.

7.8.2. Numerical analysis of the mixed finite element method for the neutron diffusion eigenproblem with heterogeneous coefficients

Participants: Patrick Ciarlet, Léandre Giret, Félix Kpadonou.

This work is done in collaboration with E. Jamelot (CEA).

We study first the convergence of the finite element approximation of the mixed diffusion equations with a source term, in the case where the solution is of low regularity. Such a situation commonly arises in the presence of three or more intersecting material components with different characteristics. Then we focus on the approximation of the associated eigenvalue problem. We prove spectral correctness for this problem in the mixed setting. These studies are carried out without, and then with a domain decomposition method. The domain decomposition method can be non-matching in the sense that the traces of the finite element spaces may not fit at the interface between subdomains. Finally, numerical experiments illustrate the accuracy of the method.

7.8.3. Localization of global norms and robust a posteriori error control for transmission problems with sign-changing coefficients

Participant: Patrick Ciarlet.

This work is done in collaboration with M. Vohralik (EPI SERENA).

We present a posteriori error analysis of diffusion problems where the diffusion tensor is not necessarily symmetric and positive definite and can in particular change its sign. We first identify the correct intrinsic error norm for such problems, covering both conforming and nonconforming approximations. It combines a dual (residual) norm together with the distance to the correct functional space. Importantly, we show the equivalence of both these quantities defined globally over the entire computational domain with the Hilbertian sums of their localizations over patches of elements. In this framework, we then design a posteriori estimators which deliver simultaneously guaranteed error upper bound, global and local error lower bounds, and robustness with respect to the (sign-changing) diffusion tensor. Robustness with respect to the approximation polynomial degree is achieved as well. The estimators are given in a unified setting covering at once conforming, nonconforming, mixed, and discontinuous Galerkin finite element discretizations in two or three space dimensions. Numerical results illustrate the theoretical developments.

7.8.4. On the convergence in H^1-norm for the fractional Laplacian

Participant: Patrick Ciarlet.

This work is done in collaboration with J.P. Borthagaray (University of Maryland).

We consider the numerical solution of the fractional Laplacian of index $s \in (1/2, 1)$ in a bounded domain Ω with homogeneous boundary conditions. Its solution a priori belongs to the fractional order Sobolev space $\widetilde{H}^s(\Omega)$. For the Dirichlet problem and under suitable assumptions on the data, it can be shown that its solution is also in $H^1(\Omega)$. In this case, if one uses the standard Lagrange finite element to discretize the problem, then both the exact and the computed solution belong to $H^1(\Omega)$. A natural question is then whether one can obtain error estimates in $H^1(\Omega)$-norm, in addition to the classical ones that can be derived in the $\widetilde{H}^s(\Omega)$ energy norm. We address this issue, and in particular we derive error estimates for the Lagrange finite element solutions on both quasi-uniform and graded meshes.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- Contract and CIFRE PhD with EDF on *the FEM-BEM coupling for soil-structure interactions*
  Participants: M. Bonnet, S. Chaillat, Z. Adnani
  Start: 11/2014. End: 02/2018. Administrator: CNRS
- Contract and CIFRE PhD with Airbus on *time-harmonic acoustic scattering in a vortical flow*
  Participants: P. Joly, J.-F. Mercier, A. Bensalah
  Start: 10/2014, End: 04/2018. Administrator: ENSTA
- Contract and CIFRE PhD with Naval Group on *modelling the fluid-structure coupling caused by a far-field underwater explosion*
  Participants: M. Bonnet, S. Chaillat, D. Mavaleix-Marchessoux
  Start: 11/2017. End: 10/2020. Administrator: CNRS
- Contract and CIFRE PhD with Naval Group on *flow noise prediction*
  Participants: J-F Mercier, S. Cotté, N. Trafny
  Start: 04/2018. End: 03/2021. Administrator: ENSTA

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

- ANR project NonlocalDD *(Non-local domain decomposition methods in electromagnetics)*
Partners: Inria Alpines, Inria POEMS, Inria Magique 3D
Start: 10/2015. End: 09/2019. Administrator: Inria
Participants of POEMS: S. Chaillat, P. Joly
Coordinator: X. Claeys (LJLL, EPI ALPINES)
- ANR project MODULATE *(Modeling lOng-perioD groUnd motions, and assessment of their effects on Large-scale infrAsTructurEs)*
Partners: ENSTA (UME), Inria POEMS, CentraleSupelec, BRGM, GDS
Start: 11/2018. End: 10/2021. Administrator: ENSTA
Participant of POEMS: S. Chaillat
Coordinator: K. Meza Fajardo (BRGM)

9.1.2. DGA

- Contracts between DGA and POEMS:
  - Contract on *inverse problems*
  Participants: L. Bourgeois
  Start: 10/2016. End: 09/2018. Administrator: ENSTA
  - Contract on *boundary element methods and high-frequency problems*
  Participants: E. Lunéville, M. Lenoir, N. Kielbasiewicz.
  Start: 10/2015. End: 2021. Administrator: ENSTA
  In partnership with F. Alouges and M. Aussal (CMAP, Ecole Polytechnique).
  - Contract on *the preconditioning of fast BEM solvers*
  Participants: S. Chaillat, F. Amlani
  Start: 10/2017. End: 12/2018. Administrator: ENSTA
- DGA provides partial funding for several PhD students:
  - A. Bera on the *design of invisible obstacles for acoustic and electromagnetic waves* (Start: 10/2016)
  - C. Beneteau on the *asymptotic analysis of time harmonic Maxwell equations in presence of metamaterials* (Start: 10/2017)
  - D. Chicaud on *domain decomposition methods for time-harmonic electromagnetic wave problems with complex media* (Start: 10/2018)

9.2. International Initiatives

9.2.1. Inria International Partners

9.2.1.1. Informal International Partners

Wilkins Aquino (Duke University, USA)

Juan Pablo Borthagaray (University of Maryland, College Park, USA)

Fioralba Cakoni (University of Rutgers, USA)

Mahadevan Ganesh (Colorado School of Mines, USA)

Camille Carvalho (UC Merced, Merced, USA)

Christophe Geuzaine (Université de Liège, Belgium)

Bojan Guzina (University of Minnesota, USA)

Marcus Grote (Universitaet Basel, Switzerland)

Jean-François Molinari (EPFL, Lausanne, Switzerland)

Sergei Nazarov (Saint-Petersburg University, Russia)

Jerónimo Rodríguez (University of Santiago de Compostela, Spain)

Adrien Semin (TU Darmstadt, Germany)

Ricardo Weder (Universidad Nacional Autonoma, Mexico)

Shravan Veerapaneni (Univ. of Michigan at Ann Arbor, USA)

Jun Zou (Chinese University of Hong Kong, HK)

9.3. International Research Visitors

9.3.1. Visits of International Scientists

- Mahadevan Ganesh (Colorado School of Mines) – July 2018
- Bojan Guzina (University of Minnesota, USA) – Summer 2018, 1 month
- Michael Weinstein (Columbia University, USA) – May-June 2018
- Fedor Bakharev (Saint Petersburg State University, Russia) – July 2018

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Advisory and management activities

- E. Lunéville is the Head of UMA (Unité de Mathématiques Appliquées) at ENSTA ParisTech.
- P. Ciarlet is coordinator of the *Mathematics & Engineering Program* of the Mathematics Hadamard Labex (LMH)

10.1.2. Scientific events organisation end selection

- S. Chaillat co-organized with X. Claeys (LJLL, EPI ALPINES) the symposium of the International Association for Boundary Element Methods (IABEM), which took place in Paris in June 2018. They were about 140 attendees.
- A.-S. Bonnet-Ben Dhia co-organized a workshop entitled *“Advanced Theoretical and Numerical Methods for waves in structured Media”* in Paris in March 2018, in the framework of the GDR Ondes. They were about 90 attendees.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- A. S. Bonnet-Ben Dhia is an associate editor of SIAP *(SIAM Journal of Applied Mathematics)*.
- M. Bonnet is an associate editor of *Engineering Analysis with Boundary Elements*, *Journal of Optimization Theory* and *Application and Journal of Integral Equations and Applications*. He is in the editorial board of *Inverse Problems* and *Computational Mechanics*.
- P. Ciarlet is an editor of ESAIM:M2AN *(Mathematical Modeling and Numerical Analysis)*.
- P. Joly is a member of the Book Series Scientific Computing of Springer Verlag.

10.1.3.2. Reviewer - Reviewing Activities

The team members regularly review papers for many international journals.

10.2. Teaching - Supervision

10.2.1. Teaching

The permanent members of POEMS are involved in the engineering program at ENSTA ParisTech, the master program *"Analyse, Modélisation et Simulation"* (AMS) and the master program *"Modélisation et Simulation en Mécanique des Structures et Systèmes Couplés"* (MS2SC) of Université Paris-Saclay.

Eliane Bécache

- *Fonctions de variable complexe*, ENSTA (1st year)
- *La méthode des éléments finis*, ENSTA (2nd year) and Master AMS (M1)
- *Analyse et approximation par éléments finis d'EDP*, ENSTA (2nd year) and Master AMS (M1)
- *Résolution des problèmes de diffraction par équations intégrales*, ENSTA (3rd year), Master AMS (M2) and Master MS2SC (M2)

Marc Bonnet

- *Problème inverses et Identification*, Master MS2SC (M2)
- Co-head of the master program *"Modélisation et Simulation en Mécanique des Structures et Systèmes Couplés"* (MS2SC) of Université Paris-Saclay

Anne-Sophie Bonnet-Ben Dhia

- *Fonctions de variable complexe*, ENSTA (1st year)
- *Théorie spectrale des opérateurs autoadjoints*, ENSTA (2nd year) and Master AMS (M1)
- *Propagation et diffraction dans les guides d'ondes*, ENSTA (3rd year) and Master AMS (M2)
- *Non Destructive Testing*, Master "Acoustical Engineering" (M2)
- *Propagation des ondes élastiques dans les solides*, Master MS2SC (M2)

Laurent Bourgeois

- *Outils élémentaires pour l'analyse des équations aux dérivées partielles*, ENSTA (1st year)
- *Fonctions de variable complexe*, ENSTA (1st year)
- *Complétion de données et identification dans les problèmes gouvernés par des équations aux dérivées partielles*, ENSTA (3rd year) and Master AMS (M2)

Stéphanie Chaillat

- *Introduction à la discrétisation des équations aux dérivées partielles*, ENSTA (1st year)
- *Fonctions de variable complexe*, ENSTA (1st year)
- *La méthode des éléments finis*, ENSTA (2nd year) and Master AMS (M1)
- *Résolution des problèmes de diffraction par équations intégrales*, ENSTA (3rd year), Master AMS (M2) and Master MS2SC (M2)
- *Equations intégrales et multipôles rapides*, Ecole doctorale MODES (Univ. Paris Est, Marne la Vallée)

Colin Chambeyron

- *Remise à niveau en maths*, Licence (1st year), Paris-Dauphine University
- *Outils mathématiques*, Licence (L1), Paris-Dauphine University
- *Analyse - Optimisation*, Licence (L1), Paris-Dauphine University
- *Algèbre linéaire*, Licence (L2), Paris-Dauphine University

Patrick Ciarlet

- *Analyse et approximation par éléments finis d'EDP*, ENSTA (2nd year) and Master AMS (M1)
- *Préformation filière ModSim*, ENSTA (3rd year)
- *Calcul scientifique parallèle*, ENSTA (3rd year) and Master AMS (M2)
- *Modèles mathématiques et leur discrétisation en électromagnétisme*, ENSTA (3rd year) and Master AMS (M2)
- Deputy head of the master program *"Analyse, Modélisation et Simulation"* of Université Paris-Saclay

Sonia Fliss

- *La méthode des éléments finis*, ENSTA (2nd year) and Master AMS (M1)
- *Introduction à la discrétisation des équations aux dérivées partielles*, ENSTA (1st year)
- *Homogénéisation périodique*, ENSTA (3rd year), Master AMS(M2), Masters ANEDP, M4S
- *Propagation des ondes dans des milieux périodiques*, ENSTA (3rd year) and Master AMS (M2)

Laure Giovangigli

- *Calcul stochastique*, ENSTA (3rd year) and Master MMMEF (M2)

Christophe Hazard

- *Outils élémentaires d'analyse pour les équations aux dérivées partielles*, ENSTA (1st year)
- *Théorie spectrale des opérateurs autoadjoints*, ENSTA (2nd year) and Master AMS (M1)

Patrick Joly

- *Introduction à la discrétisation des équations aux dérivées partielles*, ENSTA (1st year)
- *Fonctions de variable complexe*, ENSTA (1st year)
- *Analyse fonctionnelle*, ENSTA (2nd year) and Master AMS (M2)
- *Propagation des ondes dans des milieux périodiques*, ENSTA (3rd year) and Master AMS (M2)

Maryna Kachanovska

- *Fonctions de variable complexe*, ENSTA (1st year)
- *Introduction à la discrétisation des équations aux dérivées partielles*, ENSTA (1st year)
- *Systèmes Dynamiques: Analyse et Stabilité*, ENSTA (1st year)

Nicolas Kielbasiewicz

- *Programmation scientifique en C++*, ENSTA (2nd year) and Master AMS (M1)
- *Projet de simulation numérique*, ENSTA (2nd year) and Master AMS (M1)
- *Calcul scientifique parallèle*, ENSTA (3rd year) and Master AMS (M2)

Eric Lunéville

- *Introduction au calcul scientifique*, ENSTA (2nd year).
- *Programmation scientifique en C++*, ENSTA (2nd year) and Master AMS (M1)
- *Projet de simulation numérique*, ENSTA (2nd year) and Master AMS (M1)
- *Propagation et diffraction dans les guides d'ondes*, ENSTA (3rd year) and Master AMS (M2)

Jean-François Mercier

- *Outils élémentaires d'analyse pour les équations aux dérivées partielles*, ENSTA (1st year)

- *Fonctions de variable complexe*, ENSTA (1st year)
- *Théorie spectrale des opérateurs autoadjoints*, ENSTA (2nd year) and Master AMS (M1)

Axel Modave

- *La méthode des éléments finis*, ENSTA (2nd year) and Master AMS (M1)
- *Calcul scientifique à haute performance*, ENSTA (2rd year) and Master AMS (M1)
- *Calcul scientifique parallèle*, ENSTA (3rd year) and Master AMS (M2)
- *Modèles mathématiques et leur discrétisation en électromagnétisme*, ENSTA (3rd year) and Master AMS (M2)

10.2.2. Supervision

PhD: Antoine Bensalah, *"Une approche nouvelle de la modélisation mathématique et numérique en aéroacoustique par les équations de Goldstein et applications en aéronautique"*, July 2018, Patrick Joly and Jean-François Mercier

PhD: Léandre Giret, *"Analyse numérique d'une méthode de décomposition de domaine non-conforme pour les équations SPN multigroupes"*, June 2018, Patrick Ciarlet and Erell Jamelot

PhD: Zouhair Adnani, *"Modélisation numérique tridimensionnelle de l'interaction sol-structure avec prise en compte des effets de site"*, May 2018, Marc Bonnet and Stéphanie Chaillat

PhD: Arnaud Recoquillay, *"Méthodes d'échantillonnage appliquées a l'imagerie de guides d'ondes élastiques"*, January 2018, Laurent Bourgeois

PhD in progress: Antoine Bera, *"Conception de perturbations invisibles pour les ondes électromagnétiques ou acoustiques"*, October 2016, Anne-Sophie Bonnet-Ben Dhia and Lucas Chesnel

PhD in progress: Sandrine Paolantoni, *"Analyse spectrale et simulation numérique de la diffraction électromagnétique par des métamatériaux"*, October 2016, Christophe Hazard and Boris Gralak

PhD in progress: Yohanes Tjandrawidjaja, *"Modélisation de la propagation d'ondes guidées et de leur interaction avec des défauts localisés dans une plaque élastique anisotrope pour des applications en SHM"*, October 2016, Anne-Sophie Bonnet-Ben Dhia and Sonia Fliss

PhD in progress: Emile Parolin, *"Non overlapping domain decomposition methods with non local transmission conditions for electromagnetic wave propagation"*, October 2017, Patrick Joly and Xavier Claeys

PhD in progress: Clément Beneteau, *"Asymptotic analysis of time harmonic Maxwell equations in presence of metamaterials"*, October 2017, Sonia Fliss and Xavier Claeys

PhD in progress: Hajer Methenni, *"Mathematical modelling and numerical method for the simulation of ultrasound structural health monitoring of composite plates"*, October 2017, Sonia Fliss and Sébastien Impériale

PhD in progress: Damien Mavaleix, *"Modeling of the fluid-structure interaction resulting from a remote underwater explosion"*, December 2017, Marc Bonnet and Stéphanie Chaillat

PhD in progress: Damien Chicaud, *"Méthodes de décomposition de domaine pour la résolution de problèmes harmoniques d'ondes électromagnétiques en milieux complexes"*, October 2018, Patrick Ciarlet and Axel Modave

PhD in progress: Mahran Rihani, *"Équations de Maxwell en présence de méta-matériaux"*, November 2018, Anne-Sophie Bonnet-Ben Dhia and Lucas Chesnel

PhD in progress: Yacine Abourrig, *"Boundary element method for modeling electromagnetic non-destructive testing: perturbative techniques for efficient and accurate parametric studies involving multiple simulations"*, October 2017, Marc Bonnet and Edouard Demaldent

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

- M. Kachanovska is a member of the *comité scientifique* of Inria-Saclay.
- A.-S. Bonnet-Ben Dhia is a member of the *bureau du comité des équipes-projets* (BCEP).

10.3.2. Interventions

- POEMS has been involved in the exhibition *"Rencontre diffractante"*, which has been presented successively at ENSTA *(from December 2017 to February 2018)* and at Inria *(in April and May 2018)*. The inspiring material of this exhibition was made up of numerical simulations of various diffraction problems studied within by POEMS team. They served as a basis for the creation of artistic objects made by several classes of the Ecole Boulle. *Curators of the exhibition : Jérôme Perez, ENSTA, and Virginie Gannac, doctor in art, design and applied arts from Université Paris I Sorbonne, Ecole Boulle de Paris*
- POEMS has been involved in the event *"Fête de la science"* at ENSTA-ParisTech in October 2018.
- E. Bécache presented a talk entitled *"Promenade mathématique"* during the day *"Filles et mathématiques, une équation lumineuse"* (March 27, 2018 at Ecole Polytechnique, Palaiseau) and during the awards ceremony of the National Olympiad of Mathematics (May 30, 2018 at Créteil).
- E. Bécache has been involved in a meeting of young female mathematics, organized on November 17-18, 2018 at ENSTA ParisTech.

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] L. GIRET.*Numerical Analysis of a Non-Conforming Domain Decomposition for the Multigroup SPN Equations*, Université Paris-Saclay, June 2018, https://pastel.archives-ouvertes.fr/tel-01936967

[2] A. RECOQUILLAY.*Sampling methods applied to Non Destructive Testing for elastic waveguides*, Université Paris-Saclay, January 2018, https://pastel.archives-ouvertes.fr/tel-01712219

Articles in International Peer-Reviewed Journal

[3] J. ALBELLA MARTÍNEZ, S. IMPERIALE, P. JOLY, J. RODRÍGUEZ.*Solving 2D linear isotropic elastodynamics by means of scalar potentials: a new challenge for finite elements*, in "Journal of Scientific Computing", 2018 [*DOI :* 10.1007/S10915-018-0768-9], https://hal.inria.fr/hal-01803536

[4] V. BARONIAN, L. BOURGEOIS, B. CHAPUIS, A. RECOQUILLAY.*Linear Sampling Method applied to Non Destructive Testing of an elastic waveguide: theory, numerics and experiments* , in "Inverse Problems", May 2018, https://hal.inria.fr/hal-01816810

[5] A. BENSALAH, P. JOLY, J.-F. MERCIER.*Well-posedness of a generalized time-harmonic transport equation for acoustics in flow*, in "Mathematical Methods in the Applied Sciences", May 2018, vol. 41, n° 8, p. 3117 - 3137 [*DOI :* 10.1002/MMA.4805], https://hal.inria.fr/hal-01897419

[6] M. BONNET.*Inverse acoustic scattering using high-order small-inclusion expansion of misfit function*, in "Inverse Problems and Imaging ", June 2018, vol. 12, n° 4, p. 921-953 [*DOI :* 10.3934/IPI.2018039], https://hal.archives-ouvertes.fr/hal-01828087

[7] M. BONNET, R. CORNAGGIA, B. B. GUZINA.*Microstructural topological sensitivities of the second-order macroscopic model for waves in periodic media*, in "SIAM Journal on Applied Mathematics", 2018, vol. 78, n^{o} 4, p. 2057-2082 [*DOI :* 10.1137/17M1149018], https://hal.archives-ouvertes.fr/hal-01742396

[8] M. BONNET, R. CORNAGGIA, B. B. GUZINA.*Sub-wavelength sensing of bi-periodic materials using topological sensitivity of second-order homogenized model*, in "Journal of Physics: Conference Series", November 2018, vol. 1131, 012008 [*DOI :* 10.1088/1742-6596/1131/1/012008], https://hal.archives-ouvertes.fr/hal-01972083

[9] M. BONNET, E. DEMALDENT.*The eddy current model as a low-frequency, high-conductivity asymptotic form of the Maxwell transmission problem*, in "Computers & Mathematics with Applications", December 2018, https://hal.archives-ouvertes.fr/hal-01956433

[10] A.-S. BONNET-BEN DHIA, C. CARVALHO, P. CIARLET.*Mesh requirements for the finite element approximation of problems with sign-changing coefficients*, in "Numerische Mathematik", 2018, vol. 138, p. 801-838 [*DOI :* 10.1007/S00211-017-0923-5], https://hal.archives-ouvertes.fr/hal-01335153

[11] A.-S. BONNET-BEN DHIA, L. CHESNEL, S. NAZAROV.*Perfect transmission invisibility for waveguides with sound hard walls*, in "Journal de Mathématiques Pures et Appliquées", March 2018, https://hal.archives-ouvertes.fr/hal-01371163

[12] A.-S. BONNET-BEN DHIA, L. CHESNEL, V. PAGNEUX.*Trapped modes and reflectionless modes as eigenfunctions of the same spectral problem*, in "Proceedings of the Royal Society of London. Series A, Mathematical and physical sciences", May 2018, https://hal.archives-ouvertes.fr/hal-01692297

[13] A.-S. BONNET-BEN DHIA, S. FLISS, A. TONNOIR.*The Halfspace Matching Method : a new method to solve scattering problem in infinite media*, in "Journal of Computational and Applied Mathematics", July 2018, vol. 338, p. 44-68, https://hal.archives-ouvertes.fr/hal-01561339

[14] L. BOURGEOIS, A. RECOQUILLAY.*A mixed formulation of the Tikhonov regularization and its application to inverse PDE problems*, in "Modelisation Mathématique et Analyse Numérique", May 2018, https://hal.inria.fr/hal-01816804

[15] S. CHAILLAT, S. P. GROTH, A. LOSEILLE.*Metric-based anisotropic mesh adaptation for 3D acoustic boundary element methods*, in "Journal of Computational Physics", November 2018, vol. 372, p. 473 - 499 [*DOI :* 10.1016/J.JCP.2018.06.048], https://hal.archives-ouvertes.fr/hal-01895636

[16] P. CIARLET, M. VOHRALÍK.*Localization of global norms and robust a posteriori error control for transmission problems with sign-changing coefficients*, in "Modelisation Mathématique et Analyse Numérique", December 2018, vol. 52, n^{o} 5, p. 2037-2064 [*DOI :* 10.1051/M2AN/2018034], https://hal.inria.fr/hal-01148476

[17] P. JOLY, M. KACHANOVSKA, A. SEMIN.*Wave propagation in fractal trees. Mathematical and Numerical Issues*, in "Networks and Heterogeneous Media", 2018, vol. X [*DOI :* 10.3934/NHM.XX.XX.XX], https://hal.archives-ouvertes.fr/hal-01801394

[18] A. MAUREL, J.-J. MARIGO, J.-F. MERCIER, K. PHAM.*Modelling resonant arrays of the Helmholtz type in the time domain*, in "Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences", February 2018, vol. 474, n^{o} 2210 [*DOI :* 10.1098/RSPA.2017.0894], https://hal.archives-ouvertes.fr/hal-01822490

[19] J.-F. MERCIER.*Accuracy of a Low Mach Number Model for Time-Harmonic Acoustics*, in "SIAM Journal on Applied Mathematics", July 2018, vol. 78, n^{o} 4, p. 1891 - 1912 [*DOI :* 10.1137/17M113976X], https://hal.inria.fr/hal-01897611

International Conferences with Proceedings

[20] A. MODAVE, X. ANTOINE, C. GEUZAINE.*An efficient domain decomposition method with cross-point treatment for Helmholtz problems*, in "XXXIX Ibero-Latin American Congress on Computational Methods in Engineering", Compiègne/Paris, France, November 2018, https://hal.inria.fr/hal-01900309

Scientific Books (or Scientific Book chapters)

[21] A.-S. BONNET-BEN DHIA, S. FLISS, Y. TJANDRAWIDJAJA.*Numerical analysis of the Half-Space Matching method with Robin traces on a convex polygonal scatterer*, in "Maxwell's equations", De Gruyter, 2018, https://hal.inria.fr/hal-01793511

Other Publications

[22] W. AQUINO, M. BONNET.*Analysis of the error in constitutive equation approach for time-harmonic elasticity imaging*, 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01948668

[23] D. H. BAFFET, M. J. GROTE, S. IMPERIALE, M. KACHANOVSKA.*Energy Decay and Stability of a Perfectly Matched Layer For the Wave Equation*, September 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01865484

[24] L. BARATCHART, S. CHEVILLARD, J. LEBLOND, E. A. LIMA, D. PONOMAREV.*Asymptotic method for estimating magnetic moments from field measurements on a planar grid*, 2018, working paper or preprint, https://hal.inria.fr/hal-01421157

[25] M. BONNET, F. CAKONI.*Analysis of topological derivative as a tool for qualitative identification*, November 2018, https://arxiv.org/abs/1811.10538 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01929050

[26] J. P. BORTHAGARAY, P. CIARLET.*On the convergence in H^1-norm for the fractional Laplacian*, November 2018, working paper or preprint, https://hal.inria.fr/hal-01912092

[27] L. BOURGEOIS, L. CHESNEL, S. FLISS.*On well-posedness of time-harmonic problems in an unbounded strip for a thin plate model*, September 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01883560

[28] L. BOURGEOIS, D. PONOMAREV, J. DARDÉ.*An inverse obstacle problem for the wave equation in a finite time domain*, June 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01818956

[29] B. DELOURME, S. FLISS, P. JOLY, E. VASILEVSKAYA.*Trapped modes in thin and infinite ladder like domains. Part 2 : asymptotic analysis and numerical application*, June 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01822437

[30] L. FRÉROT, M. BONNET, J.-F. MOLINARI, G. ANCIAUX.*A Fourier-accelerated volume integral method for elastoplastic contact*, December 2018, https://arxiv.org/abs/1811.11558 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01963034

[31] A. MODAVE, C. GEUZAINE, X. ANTOINE.*Corner treatment for high-order local absorbing boundary conditions in high-frequency acoustic scattering*, November 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01925160

Team RANDOPT

Randomized Optimisation

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

RESEARCH CENTER
Saclay - Île-de-France

THEME
Optimization, machine learning and statistical methods

Table of contents

Team RANDOPT

Creation of the Team: 2016 December 01, updated into Project-Team: 2019 January 01

Keywords:

Computer Science and Digital Science:

A6. - Modeling, simulation and control
A6.2. - Scientific computing, Numerical Analysis & Optimization
A6.2.2. - Numerical probability
A6.2.3. - Probabilistic methods
A6.2.4. - Statistical methods
A6.2.6. - Optimization
A8.2.2. - Evolutionary algorithms
A8.9. - Performance evaluation

Other Research Topics and Application Domains:

B4. - Energy

1. Team, Visitors, External Collaborators

Research Scientists

Anne Auger [Team leader, Inria, Researcher, HDR]
Dimo Brockhoff [Inria, Researcher]
Nikolaus Hansen [Inria, Senior Researcher, HDR]

Technical Staff

Umut Batu [Inria]
Dejan Tusar [Inria, until Apr 2018]

PhD Students

Abdelbari Bouzarkouna [Inria, from Nov 2018]
Paul Dufosse [Inria, from Oct 2018]
Cheikh Saliou Toure [Inria]
Konstantinos Varelas [Thales]

Post-Doctoral Fellow

Asma Atamna [Inria, until Jun 2018]

Administrative Assistant

Natalia Alves [Inria]

2. Overall Objectives

2.1. Scientific Context

Critical problems of the 21st century like the search for highly energy efficient or even carbon-neutral, and cost-efficient systems, or the design of new molecules against extensively drug-resistant bacteria crucially rely on the resolution of challenging numerical optimization problems. Such problems typically depend on noisy experimental data or involve complex numerical simulations such that *derivatives are not useful or not available* and the function is seen as a *black-box*.

Many of those optimization problems are in essence *multiobjective*—one needs to optimize simultaneously several conflicting objectives like minimizing the cost of an energy network and maximizing its reliability—and most of the *challenging* black-box problems are *non-convex, non-smooth* and combine difficulties related to ill-conditioning, non-separability, and ruggedness (a term that characterizes functions that can be non-smooth but also noisy or multi-modal). Additionally the objective function can be expensive to evaluate, that is one function evaluation can take several minutes to hours (it can involve for instance a CFD simulation).

In this context, the use of randomness combined with proper adaptive mechanisms that particularly satisfy several invariance properties (affine invariance, invariance to monotonic transformations) has proven to be one key component for the design of robust global numerical optimization algorithms [35], [24].
The field of adaptive stochastic optimization algorithms has witnessed some important progress over the past 15 years. On the one hand, subdomains like medium-scale unconstrained optimization may be considered as “solved" (particularly, the CMA-ES algorithm, an instance of *Evolution Strategy* (ES) algorithms, stands out as state-of-the-art method) and considerably better standards have been established in the way benchmarking and experimentation are performed. On the other hand, multiobjective population-based stochastic algorithms became the method of choice to address multiobjective problems when a set of some best possible compromises is thought for. In all cases, the resulting algorithms have been naturally transferred to industry (the CMA-ES algorithm is now regularly used in companies such as Bosch, Total, ALSTOM, ...) or to other academic domains where difficult problems need to be solved such as physics, biology [38], geoscience [30], or robotics [33]).

Very recently, ES algorithms attracted quite some attention in Machine Learning with the OpenAI article *Evolution Strategies as a Scalable Alternative to Reinforcement Learning*. It is shown that the training time for difficult reinforcement learning benchmarks could be reduced from 1 day (with standard RL approaches) to 1 hour using ES [36]. [0] A few years ago, another impressive application of CMA-ES, how “Computer Sim Teaches Itself To Walk Upright” (published at the conference SIGGRAPH Asia 2013) was presented in the press in the UK.

Several of those important advances around adaptive stochastic optimization algorithms are relying to a great extent on works initiated or achieved by the founding members of RandOpt particularly related to the CMA-ES algorithm and to the Comparing Continuous Optimizer (COCO) platform.
Yet, the field of adaptive stochastic algorithms for black-box optimization is relatively young compared to the “classical optimization” field that includes convex and gradient-based optimization. For instance, the state-of-the art algorithms for unconstrained gradient based optimization like quasi-Newton methods (e.g. the BFGS method) date from the 1970s [23] while the stochastic derivative-free counterpart, CMA-ES dates from the early 2000s [25]. Consequently, in some subdomains with *important practical demands*, not even the most fundamental and basic questions are answered:

- This is the case of *constrained* optimization where one needs to find a solution $x^* \in \mathbb{R}^n$ minimizing a numerical function $\min_{x \in \mathbb{R}^n} f(x)$ while respecting a number of constraints m typically formulated as $g_i(x^*) \leq 0$ for $i = 1, ..., m$. Only recently, the fundamental requirement of linear convergence [0], as in the unconstrained case, has been clearly stated [14].

[0]The key behind such an improvement is the parallelization of the algorithm (on thousands of CPUs) that is done in such a way that the communication between the different workers is reduced to only exchanging a vector of permutation of small length (typically less than 100) containing the ranking of candidate solutions on the function to be optimized. In contrast, parallelization of backpropagation requires to exchange the gradient vector of the size of the problem (of the order of 10^6). This reduced communication time is an important factor for the important speedup.

[0]In optimization, linear convergence for an algorithm whose estimate of the optimum x^* of f at iteration t is denoted x_t, refers to a convergence where after a certain time (usually once the initialization is forgotten) the following typically holds: $\|x_{t+1} - x^*\| \leq c\|x_t - x^*\|$ where $c < 1$. This type of convergence is also called geometric. In the case of stochastic algorithms, there exist different definitions of linear convergence (depending on whether we consider the expectation of the sequence or we want a statement that holds with high probability) not strictly equivalent but that always translate the idea that the distance to the optimum at iteration $t + 1$ is a fraction of the distance to the optimum at iteration t.

- In multiobjective optimization, most of the research so far has been focusing on *how to select candidate solutions from one iteration to the next one*. The difficult question of how to *generate* effectively new solutions is not yet answered in a proper way and we know today that simply applying operators from single-objective optimization may not be effective with the current best selection strategies. As a comparison, in the single-objective case, the question of selection of candidate solutions was already solved in the 1980s and 15 more years were needed to solve the trickier question of an effective adaptive strategy to generate new solutions.
- With the current demand to solve larger and larger optimization problems (e.g. in the domain of deep learning), optimization algorithms that scale linearly (in terms of internal complexity, memory and number of function evaluations to reach an ϵ-ball around the optimum) with the problem dimension are nowadays needed. Only recently, first proposals of how to reduce the quadratic scaling of CMA-ES have been made without a clear view of what can be achieved in the best case *in practice*. These later variants apply to optimization problems with thousands of variables. The question of designing randomized algorithms capable to handle efficiently problems with one or two orders of magnitude more variables is still largely open.
- For expensive optimization, standard methods are so called Bayesian optimization (BO) algorithms often based on Gaussian processes. Commonly used examples of BO algorithms are EGO [29], SMAC [27], Spearmint [37], or TPE [17] which are implemented in different libraries. Yet, our experience with a popular method like EGO is that many important aspects to come up with a good implementation rely on insider knowledge and are not standard across implementations. Two EGO implementations can differ for example in how they perform the initial design, which bandwidth for the Gaussian kernel is used, or which strategy is taken to optimize the expected improvement.

Additionally, the **development of stochastic adaptive methods for black-box optimization has been mainly driven by heuristics and practice**—rather than a general theoretical framework—validated by intensive computational simulations. Undoubtedly, **this has been an asset as the scope of possibilities for design was not restricted by mathematical frameworks** for proving convergence. In effect, powerful stochastic adaptive algorithms for **unconstrained optimization** like the CMA-ES algorithm emerged from this approach. At the same time, naturally, **theory strongly lags behind practice**. For instance, the striking performances of CMA-ES empirically observed contrast with how little is theoretically proven on the method. This situation is clearly not satisfactory. On the one hand, theory generally lifts performance assessment from an empirical level to a conceptual one, rendering results independent from the problem instances where they have been tested. On the other hand, theory typically provides insights that change perspectives on some algorithm components. Also theoretical guarantees generally increase the trust in the reliability of a method and facilitate the task to make it accepted by wider communities.

Finally, as discussed above, the development of novel black-box algorithms strongly relies on scientific experimentation, and it is quite difficult to conduct proper and meaningful experimental analysis. This is well known for more than two decades now and summarized in this quote from Johnson in 1996

> *"the field of experimental analysis is fraught with pitfalls. In many ways, the implementation of an algorithm is the easy part. The hard part is successfully using that implementation to produce meaningful and valuable (and publishable!) research results."* [28]

Since then, quite some progress has been made to set better standards in conducting scientific experiments and benchmarking. Yet, some domains still suffer from poor benchmarking standards and from the generic problem of the lack of reproducibility of results. For instance, in multiobjective optimization, it is (still) not rare to see comparisons between algorithms made by solely visually inspecting Pareto fronts after a fixed budget. In Bayesian optimization, good performance seems often to be due to insider knowledge not always well described in papers.

2.2. Overall Objectives

In the context of black-box numerical optimization previously described, the scientific positioning of RandOpt is at the intersection between theory, algorithm design, and applications. Our vision is that the field of stochastic black-box optimization should reach the same level of maturity than gradient-based convex mathematical optimization. This entails major algorithmic developments for constrained, multi-objective and large-scale black-box optimization and major theoretical developments for analyzing current methods including the state-of-the-art CMA-ES.

The specificity in black-box optimization is that methods are intended to solve problems characterized by a *non*-property—*non*-convex, *non*-linear, *non*-smooth. This contrasts with gradient-based optimization and poses on the one hand some challenges when developing theoretical frameworks but also makes it compulsory to complement theory with empirical investigations.

Our ultimate goal is to provide software that is useful for practitioners. We see that theory is a means for this end (rather than an end in itself) and it is also our firm belief that parameter tuning is part of the designer's task.

This shapes, on the one hand, four main scientific objectives for our proposed team:

1. **develop novel theoretical frameworks** for guiding (a) the design of novel black-box methods and (b) their analysis, allowing to
2. provide **proofs ofkey features** of stochastic adaptive algorithms including the state-of-the-art method CMA-ES: linear convergence and learning of second order information.
3. develop **stochastic numerical black-box algorithms** following a **principled design** in domains with a strong practical need for much better methods namely **constrained, multiobjective, large-scale and expensive optimization**. Implement the methods such that they are easy to use. And finally, to
4. **set new standards in scientific experimentation, performance assessment and benchmarking** both for optimization on continuous or combinatorial search spaces. This should allow in particular to advance the state of **reproducibility of results of scientific papers** in optimization.

On the other hand, the above motivates our objectives with respect to dissemination and transfer:

1. develop software packages that people can directly use to solve their problems. This means having carefully thought out interfaces, generically applicable setting of parameters and termination conditions, proper treatment of numerical errors, catching properly various exceptions, etc.;
2. have direct collaborations with industrials;
3. publish our results both in applied mathematics and computer science bridging the gap between very often disjoint communities.

3. Research Program

3.1. Introduction

The lines of research we intend to pursue is organized along four axis namely developing novel theoretical framework, developing novel algorithms, setting novel standards in scientific experimentation and benchmarking and applications.

3.2. Developing Novel Theoretical Frameworks for Analyzing and Designing Adaptive Stochastic Algorithms

Stochastic black-box algorithms typically optimize **non-convex, non-smooth functions**. This is possible because the algorithms rely on weak mathematical properties of the underlying functions: the algorithms do not use the derivatives—hence the function does not need to be differentiable—and, additionally, often do not use the exact function value but instead how the objective function ranks candidate solutions (such methods are sometimes called function-value-free).(To illustrate a comparison-based update, consider an algorithm that samples λ (with λ an even integer) candidate solutions from a multivariate normal distribution. Let $x_1, ..., x_\lambda$ in $\mathbb{R}^n$ denote those λ candidate solutions at a given iteration. The solutions are evaluated on the function f to be minimized and ranked from the best to the worse:

$$f(x_{1:\lambda}) \leq ... \leq f(x_{\lambda:\lambda}) \ .$$

In the previous equation $i:\lambda$ denotes the index of the sampled solution associated to the i-th best solution. The new mean of the Gaussian vector from which new solutions will be sampled at the next iteration can be updated as

$$m \longleftarrow \frac{1}{\lambda} \sum_{i=1}^{\lambda/2} x_{i:\lambda} \ .$$

The previous update moves the mean towards the $\lambda/2$ best solutions. Yet the update is only based on the ranking of the candidate solutions such that the update is the same if f is optimized or $g \circ f$ where $g : \mathrm{Im}(f) \to \mathbb{R}$ is strictly increasing. Consequently, such algorithms are invariant with respect to strictly increasing transformations of the objective function. This entails that they are robust and their performances generalize well.)

Additionally, adaptive stochastic optimization algorithms typically have a **complex state space** which encodes the parameters of a probability distribution (e.g. mean and covariance matrix of a Gaussian vector) and other state vectors. This state-space is a **manifold**. While the algorithms are Markov chains, the complexity of the state-space makes that **standard Markov chain theory tools do not directly apply**. The same holds with tools stemming from stochastic approximation theory or Ordinary Differential Equation (ODE) theory where it is usually assumed that the underlying ODE (obtained by proper averaging and limit for learning rate to zero) has its critical points inside the search space. In contrast, in the cases we are interested in, the **critical points of the ODEs are at the boundary of the domain**.

Last, since we aim at developing theory that on the one hand allows to analyze the main properties of state-of-the-art methods and on the other hand is useful for algorithm design, we need to be careful not to use simplifications that would allow a proof to be done but would not capture the important properties of the algorithms. With that respect one tricky point is to develop **theory that accounts for invariance properties**. To face those specific challenges, we need to develop novel theoretical frameworks exploiting invariance properties and accounting for peculiar state-spaces. Those frameworks should allow researchers to analyze one of the core properties of adaptive stochastic methods, namely **linear convergence** on the widest possible class of functions.

We are planning to approach the question of linear convergence from three different complementary angles, using three different frameworks:

- the Markov chain framework where the convergence derives from the analysis of the stability of a normalized Markov chain existing on scaling-invariant functions for translation and scale-invariant algorithms [16]. This framework allows for a fine analysis where the exact convergence rate can be given as an implicit function of the invariant measure of the normalized Markov chain. Yet it requires the objective function to be scaling-invariant. The stability analysis can be particularly tricky as the Markov chain that needs to be studied writes as $\Phi_{t+1} = F(\Phi_t, W_{t+1})$ where $\{W_t : t > 0\}$ are independent identically distributed and F is typically discontinuous because the algorithms studied are comparison-based. This implies that practical tools for analyzing a standard property like irreducibility, that rely on investigating the stability of underlying deterministic control models [34], cannot be used. Additionally, the construction of a drift to prove ergodicity is particularly delicate when the state space includes a (normalized) covariance matrix as it is the case for analyzing the CMA-ES algorithm.

- The stochastic approximation or ODE framework. Those are standard techniques to prove the convergence of stochastic algorithms when an algorithm can be expressed as a stochastic approximation of the solution of a mean field ODE [19], [20], [31]. What is specific and induces difficulties for the algorithms we aim at analyzing is the **non-standard state-space** since the ODE variables correspond to the state-variables of the algorithm (e.g. $\mathbb{R}^n \times \mathbb{R}_{>0}$ for step-size adaptive algorithms, $\mathbb{R}^n \times \mathbb{R}_{>0} \times S^n_{++}$ where S^n_{++} denotes the set of positive definite matrices if a covariance matrix is additionally adapted). Consequently, the ODE can have many critical points at the boundary of its definition domain (e.g. all points corresponding to $\sigma_t = 0$ are critical points of the ODE) which is not typical. Also we aim at proving **linear convergence**, for that it is crucial that the learning rate does not decrease to zero which is non-standard in ODE method.
- The direct framework where we construct a global Lyapunov function for the original algorithm from which we deduce bounds on the hitting time to reach an ϵ-ball of the optimum. For this framework as for the ODE framework, we expect that the class of functions where we can prove linear convergence are composite of $g \circ f$ where f is differentiable and $g : \mathrm{Im}(f) \to \mathbb{R}$ is strictly increasing and that we can show convergence to a local minimum.

We expect those frameworks to be complementary in the sense that the assumptions required are different. Typically, the ODE framework should allow for proofs under the assumptions that learning rates are small enough while it is not needed for the Markov chain framework. Hence this latter framework captures better the real dynamics of the algorithm, yet under the assumption of scaling-invariance of the objective functions. Also, we expect some overlap in terms of function classes that can be studied by the different frameworks (typically convex-quadratic functions should be encompassed in the three frameworks). By studying the different frameworks in parallel, we expect to gain synergies and possibly understand what is the most promising approach for solving the holy grail question of the linear convergence of CMA-ES. We foresee for instance that similar approaches like the use of Foster-Lyapunov drift conditions are needed in all the frameworks and that intuition can be gained on how to establish the conditions from one framework to another one.

3.3. Algorithmic developments

We are planning on developing algorithms in the subdomains with strong practical demand for better methods of constrained, multiobjective, large-scale and expensive optimization.

Many of the algorithm developments, we propose, rely on the CMA-ES method. While this seems to restrict our possibilities, we want to emphasize that CMA-ES became a *family of methods* over the years that nowadays include various techniques and developments from the literature to handle non-standard optimization problems (noisy, large-scale, ...). The core idea of all CMA-ES variants—namely the mechanism to adapt a Gaussian distribution—has furthermore been shown to derive naturally from first principles with only minimal assumptions in the context of derivative-free black-box stochastic optimization [35], [24]. This is a strong justification for relying on the CMA-ES premises while new developments naturally include new techniques typically borrowed from other fields. While CMA-ES is now a full family of methods, for visibility reasons, we continue to refer often to "the CMA-ES algorithm".

3.3.1. Constrained optimization

Many (real-world) optimization problems have constraints related to technical feasibility, cost, etc. Constraints are classically handled in the black-box setting either via rejection of solutions violating the constraints—which can be quite costly and even lead to quasi-infinite loops—or by penalization with respect to the distance to the feasible domain (if this information can be extracted) or with respect to the constraint function value [21]. However, the penalization coefficient is a sensitive parameter that needs to be adapted in order to achieve a robust and general method [22]. Yet, **the question of how to handle properly constraints is largely unsolved**. The latest constraints handling for CMA-ES is an ad-hoc technique driven by many heuristics [22]. Also, it is particularly only recently that it was pointed out that **linear convergence properties should be preserved** when addressing constraint problems [14].

Promising approaches though, rely on using augmented Lagrangians [14], [15]. The augmented Lagrangian, here, is the objective function optimized by the algorithm. Yet, it depends on coefficients that are adapted online. The adaptation of those coefficients is the difficult part: the algorithm should be stable and the adaptation efficient. We believe that the theoretical frameworks developed (particularly the Markov chain framework) will be useful to understand how to design the adaptation mechanisms. Additionally, the question of invariance will also be at the core of the design of the methods: augmented Lagrangian approaches break the invariance to monotonic transformation of the objective functions, yet understanding the maximal invariance that can be achieved seems to be an important step towards understanding what adaptation rules should satisfy.

3.3.2. Large-scale Optimization

In the large-scale setting, we are interested to optimize problems with the order of 10^3 to 10^4 variables. For one to two orders of magnitude more variables, we will talk about a “very large-scale" setting.

In this context, algorithms with a quadratic scaling (internal and in terms of number of function evaluations needed to optimize the problem) cannot be afforded. In CMA-ES-type algorithms, we typically need to restrict the model of the covariance matrix to have only a linear number of parameters to learn such that the algorithms scale linearly in terms of internal complexity, memory and number of function evaluations to solve the problem. The main challenge is thus to have rich enough models for which we can efficiently design proper adaptation mechanisms. Some first large-scale variants of CMA-ES have been derived. They include the online adaptation of the complexity of the model [13], [12]. Yet so far they fail to optimize functions whose Hessian matrix has some small eigenvalues (say around 10^{-4}) some eigenvalues equal to 1 and some very large eigenvalue (say around 10^4), that is functions whose level sets have short and long axis.

Another direction, we want to pursue, is exploring the use of large-scale variants of CMA-ES to solve reinforcement learning problems [36].

Last, we are interested to investigate the very-large-scale setting. One approach consists in doing optimization in subspaces. This entails the efficient identification of relevant spaces and the restriction of the optimization to those subspaces.

3.3.3. Multiobjective Optimization

Multiobjective optimization, i.e., the simultaneous optimization of multiple objective functions, differs from single-objective optimization in particular in its optimization goal. Instead of aiming at converging to the solution with the best possible function value, in multiobjective optimization, a set of solutions [0] is sought. This set, called Pareto-set, contains all trade-off solutions in the sense of Pareto-optimality—no solution exists that is better in *all* objectives than a Pareto-optimal one. Because converging towards a set differs from converging to a single solution, it is no surprise that we might lose many good convergence properties if we directly apply search operators from single-objective methods. However, this is what has typically been done so far in the literature. Indeed, most of the research in stochastic algorithms for multiobjective optimization focused instead on the so called selection part, that decides which solutions should be kept during the optimization—a question that can be considered as solved for many years in the case of single-objective stochastic adaptive methods.

We therefore aim at rethinking search operators and adaptive mechanisms to improve existing methods. We expect that we can obtain orders of magnitude better convergence rates for certain problem types if we choose the right search operators. We typically see two angles of attack: On the one hand, we will study methods based on scalarizing functions that transform the multiobjective problem into a set of single-objective problems. Those single-objective problems can then be solved with state-of-the-art single-objective algorithms. Classical methods for multiobjective optimization fall into this category, but they all solve multiple single-objective problems subsequently (from scratch) instead of dynamically changing the scalarizing function during the search. On the other hand, we will improve on currently available population-based methods such as the first multiobjective versions of the CMA-ES. Here, research is needed on an even more fundamental level such as

[0] Often, this set forms a manifold of dimension one smaller than the number of objectives.

trying to understand success probabilities observed during an optimization run or how we can introduce non-elitist selection (the state of the art in single-objective stochastic adaptive algorithms) to increase robustness regarding noisy evaluations or multi-modality. The challenge here, compared to single-objective algorithms, is that the quality of a solution is not anymore independent from other sampled solutions, but can potentially depend on all known solutions (in the case of three or more objective functions), resulting in a more noisy evaluation as the relatively simple function-value-based ranking within single-objective optimizers.

3.3.4. Expensive Optimization

In the so-called expensive optimization scenario, a single function evaluation might take several minutes or even hours in a practical setting. Hence, the available budget in terms of number of function evaluation calls to find a solution is very limited in practice. To tackle such expensive optimization problems, it is needed to exploit the first few function evaluations in the best way. To this end, typical methods couple the learning of a surrogate (or meta-model) of the expensive objective function with traditional optimization algorithms.

In the context of expensive optimization and CMA-ES, which usually shows its full potential when the number n of variables is not too small (say larger than 3) and if the number of available function evaluations is about $100n$ or larger, several research directions emerge. The two main possibilities to integrate meta-models into the search with CMA-ES type algorithms are (i) the successive injection of the minimum of a learned meta-model at each time step into the learning of CMA-ES's covariance matrix and (ii) the use of a meta-model to predict the internal ranking of solutions. While for the latter, first results exist, the former idea is entirely unexplored for now. In both cases, a fundamental question is which type of meta-model (linear, quadratic, Gaussian Process, ...) is the best choice for a given number of function evaluations (as low as one or two function evaluations) and at which time the type of the meta-model shall be switched.

3.4. Setting novel standards in scientific experimentation and benchmarking

Numerical experimentation is needed as a complement to theory to test novel ideas, hypotheses, the stability of an algorithm, and/or to obtain quantitative estimates. Optimally, theory and experimentation go hand in hand, jointly guiding the understanding of the mechanisms underlying optimization algorithms. Though performing numerical experimentation on optimization algorithms is crucial and a common task, it is non-trivial and easy to fall in (common) pitfalls as stated by J. N. Hooker in his seminal paper [26].

In the RandOpt team we aim at raising the standards for both scientific experimentation and benchmarking.

On the experimentation aspect, we are convinced that there is common ground over how scientific experimentation should be done across many (sub-)domains of optimization, in particular with respect to the visualization of results, testing extreme scenarios (parameter settings, initial conditions, etc.), how to conduct understandable and small experiments, how to account for invariance properties, performing scaling up experiments and so forth. We therefore want to formalize and generalize these ideas in order to make them known to the entire optimization community with the final aim that they become standards for experimental research.

Extensive numerical benchmarking, on the other hand, is a compulsory task for evaluating and comparing the performance of algorithms. It puts algorithms to a standardized test and allows to make recommendations which algorithms should be used preferably in practice. To ease this part of optimization research, we have been developing the Comparing Continuous Optimizers platform (COCO) since 2007 which allows to automatize the tedious task of benchmarking. It is a game changer in the sense that the freed time can now be spent on the scientific part of algorithm design (instead of implementing the experiments, visualization, statistical tests, etc.) and it opened novel perspectives in algorithm testing. COCO implements a thorough, well-documented methodology that is based on the above mentioned general principles for scientific experimentation.

Also due to the freely available data from 200+ algorithms benchmarked with the platform, COCO became a quasi-standard for single-objective, noiseless optimization benchmarking. It is therefore natural to extend the reach of COCO towards other subdomains (particularly constrained optimization, many-objective optimization) which can benefit greatly from an automated benchmarking methodology and standardized tests without (much) effort. This entails particularly the design of novel test suites and rethinking the methodology for measuring performance and more generally evaluating the algorithms. Particularly challenging is the design of scalable non-trivial testbeds for constrained optimization where one can still control where the solutions lies. Other optimization problem types, we are targeting are expensive problems (and the Bayesian optimization community in particular, see our AESOP project), optimization problems in machine learning (for example parameter tuning in reinforcement learning), and the collection of real-world problems from industry.

Another aspect of our future research on benchmarking is to investigate the large amounts of benchmarking data, we collected with COCO during the years. Extracting information about the influence of algorithms on the best performing portfolio, clustering algorithms of similar performance, or the automated detection of anomalies in terms of good/bad behavior of algorithms on a subset of the functions or dimensions are some of the ideas here.

Last, we want to expand the focus of COCO from automatized (large) benchmarking experiments towards everyday experimentation, for example by allowing the user to visually investigate algorithm internals on the fly or by simplifying the set up of algorithm parameter influence studies.

4. Application Domains

4.1. Application Domains

Applications of black-box algorithms occur in various domains. Industry but also researchers in other academic domains have therefore a great need to apply black-box algorithms on a daily basis. We see this as a great source of motivation to design better methods. Applications not only allow us to backup our methods and understand what are the relevant features to solve a real-world problem but also help identify novel difficulties or set priorities in terms of algorithm design.

Asides from the two applications to Machine Learning that we detail below, we however do not target a specific application domain and we are interested in possible black-box applications stemming from various origins. This is for us intrinsic to the nature of the methods we develop that are general purpose algorithms. Hence our strategy with respect to applications can be seen as opportunistic and our main selection criteria when approached by colleagues who want to develop a collaboration around an application is whether we judge the application interesting: that is the application brings new challenges and/or gives us the opportunity to work on topics we already intended to work on.

The three concrete applications related to industrial collaborations we are currently dealing with are:

- With EDF R&D through the design and placement of bi-facial photovoltaic panel for the postdoc of Asma Atamna funded by the PGMO project.
- With Thales for the thesis of Konstantinos Varelas (DGA-CIFRE thesis) related to the design of radars (shape optimization of the wave form). This thesis investigates the development of large-scale variants of CMA-ES.
- With Storengy, a subsidiary of Engie specialized in gas storage for the thesis of Cheikh Touré. Different multi-objective applications are considered in this context but the primary motivation of Storengy is to get at their disposal a better multi-objective variant of CMA-ES which is the main objective of the developments within the thesis.

Additionally, there are two specific types of applications stemming from Machine Learning we would like to focus on: problems with non-differentiable loss that can occur in reinforcement learning and hyperparameter tuning problems. For the first class of problems the motivation comes from the paper [36] where different reinforcement learning problems are addressed and the weights of neural networks are adjusted using evolution strategies. Those problems are large-scale (in [36] up to 10^6 weights are adjusted), and the large-scale variants of CMA-ES we want to investigate might be relevant in this case. For the second class of problems (hyperparameter tuning problems), standard approches to handle those problems are Bayesian optimization algorithms but despite the tremendous effort for developing Bayesian optimization techniques and having implementations of Bayesian optimization algorithms within libraries, pure random search is still often used for training neural networks. One reason is that pure random search is intrinsically parallel [18]. This suggests that methods like CMA-ES—that are also intrinsically parallel—can be also advantageously used for hyperparameter tuning: this was demonstrated to tune deep neural networks in [32]. One limitation though of the CMA-ES algorithm is that it cannot deal with categorical/integer and continuous variables at the same time. This motivates us to investigate the development of CMA-ES variants that are able to deal with mixed variables.

When dealing with single applications, the results observed are difficult to generalize: typically not many methods are tested on a single application as tests are often time consuming and performed in restrictive settings. Yet, if one circumvent the problem of confidentiality of data and of criticality for companies to publish their applications, real-world problems could become benchmarks as any other analytical function. This would allow to test wider ranges of methods on the problems and to find out whether analytical benchmarks properly capture real-world problem difficulties. We will thus seek to incorporate real-world problems within the COCO platform. This is a recurrent demand by researchers in optimization. As far as confidentiality of data are concerned, our preliminary discussions with industrials allow us to be optimistic that we can convince industrials to propose real-world problems with anonymized (and uncritical) data that still capture the essence of the underlying real-world problem.

5. Highlights of the Year

5.1. Highlights of the Year

A. Auger appointed general chair of the ACM GECCO 2019 conference (GECCO being the largest most selective conference in EC)

6. New Software and Platforms

6.1. COCO

COmparing Continuous Optimizers
KEYWORDS: Benchmarking - Numerical optimization - Black-box optimization - Stochastic optimization
SCIENTIFIC DESCRIPTION: COmparing Continuous Optimisers (COCO) is a tool for benchmarking algorithms for black-box optimisation. COCO facilitates systematic experimentation in the field of continuous optimization. COCO provides: (1) an experimental framework for testing the algorithms, (2) post-processing facilities for generating publication quality figures and tables, (3) LaTeX templates for scientific articles and HTML overview pages which present the figures and tables.

The COCO software is composed of two parts: (i) an interface available in different programming languages (C/C++, Java, Matlab/Octave, Python, external support for R) which allows to run and log experiments on several function test suites (unbounded noisy and noiseless single-objective functions, unbounded noiseless multiobjective problems, constrained problems) are provided (ii) a Python tool for generating figures and tables that can be looked at in every web browser and that can be used in the provided LaTeX templates to write scientific papers.

FUNCTIONAL DESCRIPTION: The Coco platform aims at supporting the numerical benchmarking of blackbox optimization algorithms in continuous domains. Benchmarking is a vital part of algorithm engineering and a necessary path to recommend algorithms for practical applications. The Coco platform releases algorithm developers and practitioners alike from (re-)writing test functions, logging, and plotting facilities by providing an easy-to-handle interface in several programming languages. The Coco platform has been developed since 2007 and has been used extensively within the “Blackbox Optimization Benchmarking (BBOB)” workshop series since 2009. Overall, 160+ algorithms and algorithm variants by contributors from all over the world have been benchmarked on the platform’s three supported test suites so far. The most recent extension towards bi-objective problems has been used for the BBOB-2016 workshop at GECCO and we are currently developing new test suites around large-scale and constrained optimization.

- Participants: Anne Auger, Asma Atamna, Dejan Tusar, Dimo Brockhoff, Marc Schoenauer, Nikolaus Hansen, Ouassim Ait Elhara, Raymond Ros, Tea Tusar, Thanh-Do Tran and Umut Batu
- Partners: TU Dortmund University - Charles University Prague - Jozef Stefan Institute (JSI)
- Contact: Dimo Brockhoff
- URL: https://github.com/numbbo/coco

6.2. CMA-ES

Covariance Matrix Adaptation Evolution Strategy
KEYWORDS: Numerical optimization - Black-box optimization - Stochastic optimization
SCIENTIFIC DESCRIPTION: The CMA-ES is considered as state-of-the-art in evolutionary computation and has been adopted as one of the standard tools for continuous optimisation in many (probably hundreds of) research labs and industrial environments around the world. The CMA-ES is typically applied to unconstrained or bounded constraint optimization problems, and search space dimensions between three and a hundred. The method should be applied, if derivative based methods, e.g. quasi-Newton BFGS or conjugate gradient, (supposedly) fail due to a rugged search landscape (e.g. discontinuities, sharp bends or ridges, noise, local optima, outliers). If second order derivative based methods are successful, they are usually faster than the CMA-ES: on purely convex-quadratic functions, f(x)=xTHx, BFGS (Matlabs function fminunc) is typically faster by a factor of about ten (in terms of number of objective function evaluations needed to reach a target function value, assuming that gradients are not available). On the most simple quadratic function f(x)=||x||2=xTx BFGS is faster by a factor of about 30.
FUNCTIONAL DESCRIPTION: The CMA-ES is an evolutionary algorithm for difficult non-linear non-convex black-box optimisation problems in continuous domain.

- Participant: Nikolaus Hansen
- Contact: Nikolaus Hansen
- URL: http://cma.gforge.inria.fr/cmaes_sourcecode_page.html

6.3. Platforms

6.3.1. New developments around COCO

There were two public releases of the COCO software this year including quite some new features that have also been used for the Blackbox Optimization Benchmarking workshop (BBOB) which was held in Kyoto, Japan during GECCO-2019.

The most important new features are updated, streamlined plots, a Python 3 compatible postprocessing module with a corresponding restructuring of the postprocessing code, the support for zip files in the postprocessing, a simplified example experiment script for beginners and a non-anytime example experiment for benchmarking budget-dependent algorithms, improved coverage of the continuous integration testing via CircleCI and AppVeyor, and finally and most-important from a practical perspective an archive with automatized download from all 200+ algorithm data sets available in the COCO data archive. Of these, 17 algorithm data sets have been made newly available in 2018 with four scientific papers being presented at the BBOB-2019 workshop.

In the background, there have been additional (preparational) activities, in particular due to the two Inria ADT projects “COCOpysuites” and “COCOpost”. The “COCOpysuites” project aimed at a rewriting of the experimental part of COCO in python to allow for an easier development, testing, and implementation of new test suites. The “COCOpost” project aimed at a complete rewrite of the python postprocessing with a focus on new, interactive plots and a clearer structure for improved maintenance. In addition, new test suites have been developed and implemented for large-scale, constrained, multiobjective, and mixed-integer optimization. All those extensions will be made available step-by-step to the scientific community after proper alpha- and beta-testing in the coming planned releases.

6.3.2. Developments within the CMA-ES library

The `pycma` library has not seen major changes, but overall 39 commits pushed for maintenance, bug-fixes and smaller improvements (roughly 1000 lines of code). An as of yet unpublished development has been the modularization of the data logger. A surrogate fitness model module with 969 lines of code has been developed and is already operative but has also not yet been released.

7. New Results

7.1. Analysis of adaptive Stochastic Optimizers

7.1.1. New ODE method for proving the geometric convergence of adaptive Stochastic Optimizers

The ODE method is a standard technique to analyze the convergence of stochastic algorithms defined as a stochastic approximation of an ODE. In a nutshell, the convergence of the algorithms derives from the stability of the ODE and the control of the error between the solution of the ODE and the trajectory of the stochastic algorithm. We have been developing a new ODE method to be able to prove the geometric convergence of stochastic approximation algorithms that derive from the family of adaptive stochastic optimization algorithms. Standard theory did not apply in this context as the state variable adapted typically converge to the boundary of the state-space domain where an infinite number of points are equilibrium points for the ODE [7].

7.1.2. Convergence and convergence rate analysis of the (1+1)-ES with one-fifth success rule

When analyzing adaptive stochastic optimizers, one is typically interested to prove the linear convergence and investigate the dependency of the convergence rate with respect to the dimension. We have greatly simplified the analysis of the convergence and convergence rate of the (1+1)-ES with one-fifth success rule on the sphere function. We have shown that the analysis derives from applying a simple "drift" theorem and consequently shown a hitting time to reach an ϵ-ball of the optimum of $\Theta(\frac{1}{d}\log(1/\epsilon))$ akin to linear convergence with a convergence rate scaling linearly with the dimension [4].

7.1.3. Quality-gain Analysis on Convex-quadratic functions

We have analyzed the expected function value decrease (related to the convergence rate) of Evolution Strategies with weighed recombination on convex-quadratic functions. We have derive different bounds and limit expression that allow to derive optimal recombination weights and the optimal step-size, and found that the optimal recombination weights are independent of the Hessian of the objective function. We have moreover shown the dependencies of the optimal parameters in the dimension and population size [1].

7.2. Benchmarking Methodology

7.2.1. Single-Objective Benchmarking

Benchmarking optimization algorithms seems trivial at first sight but is quite involved in practice and little decisions on the experimental setup can have a large effect on the displayed algorithm performance.

We have investigated some of these effects in the context of the Black-Box Optimization Benchmarking (bbob) test suite of the COCO platform and the well-known quasi-Newton BFGS algorithm, default in MATLAB's fminunc and in Python's scipy.optimize module [5]. We realized in particular that the instance instantiation in the COCO platform has little impact while the initial search point has a larger one. The largest performance differences, however, stem from implementation details that are typically not documented and not exposed to the user via internal algorithm parameters. For example is the MATLAB implementation of the BFGS algorithm significantly worse than the Python implementation and the MATLAB 2017 version is worse than the MATLAB 2009 implementation.

Additionally, Nikolaus Hansen gave a hands-on tutorial on good benchmarking practice at the GECCO-2018 conference in Kyoto [11].

7.2.2. Multi-Objective Benchmarking

In terms of multiobjective benchmarking, our contributions are two-fold. Firstly, we wrote a scientific article on the scientific methodology for defining our new multiobjective benchmark suite. In [10] we introduce two new bi-objective test suites on the basis of the above mentioned, well-known 24 bbob test functions and propose a generic test suite generator for an arbitrary number of objectives. The former are implemented in our COCO platform and extensively documented in terms of search and objective space plots for each function.

Secondly, we realized with the proposal of the biobjective bbob test suites that there is a need for more theoretical analyses of simple test functions that still test for practical challenges such as ill-conditioning or search space rotations. In our upcoming EMO conference paper [3] we therefore characterize theoretically Pareto sets and Pareto fronts of combinations of two convex quadratic functions with arbitrary search space dimension. Based on this theoretical analysis, we suggest a wide set of new biobjective test functions.

7.3. Large-scale Optimization

We have been studying different large-scale variants of the CMA-ES algorithm and tested them thoroughly empirically on a set of scalable large-scale testbed. The study includes comparison with the large-scale quasi-Newton algorithm, namely L-BFGS [6].

7.4. Constrained Optimization

In the context of constrained optimization, A. Atamna studied invariance properties of Augmented Lagrangian approaches and showed the relation between invariance to strictly increasing affine transformations of the objective function and the scaling of the constraints and linear convergence [2]. Progress were made towards a methodology to define scalable constrained problems with control optimum and the implementation of a new constrained testbed within the COCO platform. In her internship, E. Marescaux studied the connection between augmented Lagrangian approaches and a previously proposed adaptive constrained handling mechanism. She also studied a new idea to turn a constrained problem into an unconstrained one.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- Contract with the company Storengy partially funding the PhD thesis of Cheikh Touré (2017 - 2020)
- Contract with Thales in the context of the CIFRE PhD thesis of Konstantinos Varelas (2017 - 2020)

9. Partnerships and Cooperations

9.1. Regional Initiatives

- PGMO/FMJH project “AESOP: Algorithms for Expensive Simulation-Based Optimization”, 7kEUR, 2017–2019

9.2. National Initiatives

9.2.1. ANR

- ANR project “Big Multiobjective Optimization (BigMO)”, Dimo Brockhoff participates in this project through the Inria team BONUS in Lille (2017–2020)

9.3. International Research Visitors

9.3.1. Visitors to RandOpt

- Filip Matzner, October 15–19 and December 10-14, 2018

9.3.2. Internships

- Eugenie Marescaux, March–July 2018
- Xudong Zhang, March–August 2018

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

- Anne Auger is the General Chair of the forthcoming ACM GECCO 2019 conference, Prague CZ (largest and most prestigious conference in the Evolutionary Computation domain).

10.1.2. Member of the Organizing Committees

- Anne Auger, Dimo Brockhoff and Nikolaus Hansen, co-organizer of the ACM-GECCO-2018 workshop on Black Box Optimization Benchmarking, together with Julien Bect, Rodolphe Le Riche, Victor Picheny, and Tea Tušar
- Anne Auger, Dimo Brockhoff, Nikolaus Hansen, and Konstantinos Varelas, co-organizer of the ACM-GECCO-2019 workshop on Black Box Optimization Benchmarking, together with Tea Tušar
- Dimo Brockhoff: co-organization of the Lorentz Center workshop on Many-Criteria Optimization (MACODA), September 2019, with Boris Naujoks, Michael Emmerich, and Robin Purshouse

10.1.3. Scientific Events Selection

10.1.3.1. Chair of Conference Program Committees

- Anne Auger has been program chair of the PPSN 2018 conference, Coimbra, Portugal.
- Anne Auger was Theory track chair for the ACM GECCO 2018 conference, Kyoto, Japan.
- Nikolaus Hansen was ENUM track chair for the ACM GECCO 2018 conference, Kyoto, Japan.

10.1.3.2. Member of the Conference Program Committees

- Dimo Brockhoff and Nikolaus Hansen were members of the program committee of the PPSN 2018 conference.

10.1.3.3. Reviewer

- Dimo Brockhoff: GECCO’2018, GECCO’2018 student workshop, PPSN’2018, LeGO 2018, EMO’2019, and FOGA’2019

10.1.4. Journal

10.1.4.1. Member of the Editorial Boards

- Anne Auger and Nikolaus Hansen members of the editorial board of the Evolutionary Computation journal.

10.1.4.2. Reviewer - Reviewing Activities

- The three permanent members are frequent reviewers for the main two journals on Evolutionary Computation: IEEE transaction on Evolutionary Computation, Evolutionary Computation.
- Anne Auger is guest editor of Algorithmica special issue of papers selected from GECCO theory tracks 2018
- Anne Auger is guest editor of IEEE Transactions on Evolutionary Computation special issue on Theoretical Foundations of Evolutionary Computation 2018/2019

10.1.5. Invited Talks

- Dimo Brockhoff: “Benchmarking multiobjective optimizers: An algorithmic jam session of recent results”, December 2018, Centre for Informatics and Systems, University of Coimbra, Portugal

10.1.6. Leadership within the Scientific Community

- Anne Auger, Elected Member of the ACM-SIGEVO executive board
- Dimo Brockhoff, member of the International Advisory Committee for EMO-2019 in East Lansing, USA

10.1.7. Scientific Expertise

- Anne Auger scientific expert for an audit on Artificial Intelligence of a large French industrial consortium.
- Dimo Brockhoff, external reviewer for the Luxembourg National Research Fund (FNR) in the CORE 2018 call

10.1.8. Research Administration

- Anne Auger, member of the conseil de laboratoire of the CMAP, Ecole Polytechnique.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- Master: Anne Auger, Course on Derivative-free Optimization of AMS and Optimization masters, 22.5 hours M2, Paris-Saclay University
- Master: Anne Auger, exercises for the courses Introduction to Machine Learning (MAP 534), M2 X/HEC and Advanced Machine Learning (MAP 541), 30 hours, Ecole Polytechnique, France
- Master: Anne Auger, exercises for the courses Advanced Machine Learning (MAP 541), M2 X/HEC and Advanced Machine Learning (MAP 541), 30 hours, Ecole Polytechnique, France
- Master: Anne Auger, 2nd year students of Ecole Polytechnique, Datacamp (solving practical machine learning problems) MAP 583, 20 hours, Ecole Polytechnique, France
- Master: Anne Auger, “Advanced Optimization”, 9h ETD, M2, Université Paris-Sud, France (joint course with D. Brockhoff)
- Master: Dimo Brockhoff, “Introduction to Optimization”, 31.5h ETD, M2, Université Paris-Sud, France
- Master: Dimo Brockhoff, “Advanced Optimization”, 22.5h ETD, M2, Université Paris-Sud, France (joint course with A. Auger)

10.2.2. Tutorials

- Dimo Brockhoff gave the tutorial “Evolutionary Multiobjective Optimization” at the GECCO-2018 conference in Kyoto, Japan [9]
- Nikolaus Hansen gave the tutorials “CMA-ES and Advanced Adaptation Mechanisms” [8] and “A Practical Guide to Experimentation (and Benchmarking)” [11] at the GECCO-2018 conference in Kyoto, Japan.

10.2.3. Supervision

- PhD in progress: Konstantinos Varelas, “Large-Scale Optimization, CMA-ES and Radar Applications”, Dec. 2017, Anne Auger and Dimo Brockhoff
- PhD in progress: Cheikh Touré, “Linearly Convergent Multi-objective Stochastic Optimizers”, Dec. 2017, Anne Auger and Dimo Brockhoff
- PhD in progress: Paul Dufossé, “Constrained Optimization and Radar Applications”, Oct. 2018, Nikolaus Hansen
- PhD in progress: Marie-Ange Dahito, “Mixed-Integer Blackbox Optimization for Multiobjective Problems in the Automotive Industry”, Jan 2019, Dimo Brockhoff and Nikolaus Hansen

10.2.4. Juries

- Anne Auger, member of Hiring committee for assistant professor positions for Toulouse 1 Capitole University, MCF 0225, section CNU 27
- Anne Auger, PhD jury of Asmaa Ghoumari, defense in Dec. 2018
- Dimo Brockhoff, PhD jury of Lukas Bajer, Charles University Prague, Czech Republic, defense in June 2018
- Dimo Brockhoff, PhD jury of Andreia P. Guerrero, University of Coimbra, Portugal, defene in December 2018

10.3. Popularization

10.3.1. Interventions

- National events: Cheikh Touré made an intervention on Poker for "college" students at the Fête de la Science (explaining relation between the game and probability)

11. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] Y. Akimoto, A. Auger, N. Hansen. *Quality Gain Analysis of the Weighted Recombination Evolution Strategy on General Convex Quadratic Functions*, in "Theoretical Computer Science", 2018 [*DOI :* 10.1016/j.tcs.2018.05.015], https://hal.inria.fr/hal-01662568

[2] A. Atamna, A. Auger, N. Hansen. *On Invariance and Linear Convergence of Evolution Strategies with Augmented Lagrangian Constraint Handling*, in "Theoretical Computer Science", November 2018, https://hal.inria.fr/hal-01660728

Invited Conferences

[3] C. Touré, A. Auger, D. Brockhoff, N. Hansen. *On Bi-Objective convex-quadratic problems*, in "10th International Conference on Evolutionary Multi-Criterion Optimization", East Lansing, Michigan, United States, March 2019, https://arxiv.org/abs/1812.00289 , https://hal.inria.fr/hal-01942159

International Conferences with Proceedings

[4] Y. AKIMOTO, A. AUGER, T. GLASMACHERS.*Drift Theory in Continuous Search Spaces: Expected Hitting Time of the (1+1)-ES with 1/5 Success Rule*, in "Proceedings of the GECCO 2018 Conference", Kyoto, Japan, 2018, https://arxiv.org/abs/1802.03209 , https://hal.inria.fr/hal-01778116

[5] A. BLELLY, M. FELIPE-GOMES, A. AUGER, D. BROCKHOFF.*Stopping Criteria, Initialization, and Implementations of BFGS and their Effect on the BBOB Test Suite*, in "GECCO '18 Companion", Kyoto, Japan, July 2018 [*DOI :* 10.1145/3205651.3208303], https://hal.inria.fr/hal-01811588

[6] K. VARELAS, A. AUGER, D. BROCKHOFF, N. HANSEN, O. A. ELHARA, Y. SEMET, R. KASSAB, F. BARBARESCO.*A Comparative Study of Large-scale Variants of CMA-ES*, in "PPSN XV 2018 - 15th International Conference on Parallel Problem Solving from Nature", Coimbra, Portugal, LNCS, September 2018, vol. 11101, p. 3-15 [*DOI :* 10.1007/978-3-319-99253-2_1], https://hal.inria.fr/hal-01881454

Other Publications

[7] Y. AKIMOTO, A. AUGER, N. HANSEN.*An ODE Method to Prove the Geometric Convergence of Adaptive Stochastic Algorithms*, November 2018, https://arxiv.org/abs/1811.06703 - working paper or preprint, https://hal.inria.fr/hal-01926472

[8] Y. AKIMOTO, N. HANSEN.*CMA-ES and Advanced Adaptation Mechanisms*, July 2018, GECCO '18 Companion: Proceedings of the Genetic and Evolutionary Computation Conference Companion, https://hal.inria.fr/hal-01959479

[9] D. BROCKHOFF.*GECCO 2018 tutorial on evolutionary multiobjective optimization*, July 2018, GECCO '18 Companion: Proceedings of the Genetic and Evolutionary Computation Conference Companion, https://hal.inria.fr/hal-01943586

[10] D. BROCKHOFF, T. TUSAR, A. AUGER, N. HANSEN.*Using Well-Understood Single-Objective Functions in Multiobjective Black-Box Optimization Test Suites*, January 2019, https://arxiv.org/abs/1604.00359 - ArXiv e-prints, arXiv:1604.00359, https://hal.inria.fr/hal-01296987

[11] N. HANSEN.*A Practical Guide to Experimentation (and Benchmarking)*, July 2018, GECCO '18 Companion: Proceedings of the Genetic and Evolutionary Computation Conference Companion, https://hal.inria.fr/hal-01959453

References in notes

[12] Y. AKIMOTO, N. HANSEN.*Online model selection for restricted covariance matrix adaptation*, in "International Conference on Parallel Problem Solving from Nature", Springer, 2016, p. 3–13

[13] Y. AKIMOTO, N. HANSEN.*Projection-based restricted covariance matrix adaptation for high dimension*, in "Proceedings of the 2016 on Genetic and Evolutionary Computation Conference", ACM, 2016, p. 197–204

[14] D. V. ARNOLD, J. PORTER.*Towards au Augmented Lagrangian Constraint Handling Approach for the* $(1+1)$*-ES*, in "Genetic and Evolutionary Computation Conference", ACM Press, 2015, p. 249-256

[15] A. ATAMNA, A. AUGER, N. HANSEN.*Linearly Convergent Evolution Strategies via Augmented Lagrangian Constraint Handling*, in "Foundation of Genetic Algorithms (FOGA)", 2017

[16] A. AUGER, N. HANSEN.*Linear Convergence of Comparison-based Step-size Adaptive Randomized Search via Stability of Markov Chains*, in "SIAM Journal on Optimization", 2016, vol. 26, n^{o} 3, p. 1589-1624

[17] J. BERGSTRA, R. BARDENET, Y. BENGIO, B. KÉGL.*Algorithms for Hyper-Parameter Optimization*, in "Neural Information Processing Systems (NIPS 2011)", 2011, https://hal.inria.fr/hal-00642998/file/draft1.pdf

[18] J. BERGSTRA, Y. BENGIO.*Random search for hyper-parameter optimization*, in "Journal of Machine Learning Research", 2012, vol. 13, p. 281–305

[19] V. S. BORKAR.*Stochastic approximation: a dynamical systems viewpoint*, 2008, Cambridge University Press

[20] V. BORKAR, S. MEYN.*The O.D.E. Method for Convergence of Stochastic Approximation and Reinforcement Learning*, in "SIAM Journal on Control and Optimization", January 2000, vol. 38, n^{o} 2

[21] C. A. COELLO COELLO.*Constraint-handling techniques used with evolutionary algorithms*, in "Proceedings of the 2008 Genetic and Evolutionary Computation Conference", ACM, 2008, p. 2445–2466

[22] G. COLLANGE, S. REYNAUD, N. HANSEN.*Covariance Matrix Adaptation Evolution Strategy for Multidisciplinary Optimization of Expendable Launcher Families*, in "13th AIAA/ISSMO Multidisciplinary Analysis Optimization Conference, Proceedings", 2010

[23] J. E. DENNIS, R. B. SCHNABEL.*Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1983

[24] N. HANSEN, A. AUGER.*Principled design of continuous stochastic search: From theory to practice*, in "Theory and principled methods for the design of metaheuristics", Springer, 2014, p. 145–180

[25] N. HANSEN, A. OSTERMEIER.*Completely Derandomized Self-Adaptation in Evolution Strategies*, in "Evolutionary Computation", 2001, vol. 9, n^{o} 2, p. 159–195

[26] J. N. HOOKER.*Testing heuristics: We have it all wrong*, in "Journal of heuristics", 1995, vol. 1, n^{o} 1, p. 33–42

[27] F. HUTTER, H. HOOS, K. LEYTON-BROWN.*An Evaluation of Sequential Model-based Optimization for Expensive Blackbox Functions*, in "GECCO (Companion) 2013", ACM, 2013, p. 1209–1216

[28] D. S. JOHNSON.*A theoretician's guide to the experimental analysis of algorithms*, in "Data structures, near neighbor searches, and methodology: fifth and sixth DIMACS implementation challenges", 2002, vol. 59, p. 215–250

[29] D. R. JONES, M. SCHONLAU, W. J. WELCH.*Efficient global optimization of expensive black-box functions*, in "Journal of Global optimization", 1998, vol. 13, n^{o} 4, p. 455–492

[30] I. KRIEST, V. SAUERLAND, S. KHATIWALA, A. SRIVASTAV, A. OSCHLIES.*Calibrating a global three-dimensional biogeochemical ocean model (MOPS-1.0)*, in "Geoscientific Model Development", 2017, vol. 10, n^{o} 1, 127

[31] H. J. KUSHNER, G. YIN.*Stochastic approximation and recursive algorithms and applications*, Applications of mathematics, Springer, New York, 2003, http://opac.inria.fr/record=b1099801

[32] I. LOSHCHILOV, F. HUTTER.*CMA-ES for hyperparameter optimization of deep neural networks*, in "arXiv preprint arXiv:1604.07269", 2016

[33] P. MACALPINE, S. BARRETT, D. URIELI, V. VU, P. STONE.*Design and Optimization of an Omnidirectional Humanoid Walk: A Winning Approach at the RoboCup 2011 3D Simulation Competition*, in "Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence (AAAI)", July 2012

[34] S. MEYN, R. TWEEDIE.*Markov Chains and Stochastic Stability*, Springer-Verlag, New York, 1993

[35] Y. OLLIVIER, L. ARNOLD, A. AUGER, N. HANSEN.*Information-geometric optimization algorithms: A unifying picture via invariance principles*, in "Journal Of Machine Learning Research", 2016, accepted

[36] T. SALIMANS, J. HO, X. CHEN, I. SUTSKEVER.*Evolution strategies as a scalable alternative to reinforcement learning*, in "arXiv preprint arXiv:1703.03864", 2017

[37] J. SNOEK, H. LAROCHELLE, R. P. ADAMS.*Practical bayesian optimization of machine learning algorithms*, in "Neural Information Processing Systems (NIPS 2012)", 2012, p. 2951–2959

[38] J. UHLENDORF, A. MIERMONT, T. DELAVEAU, G. CHARVIN, F. FAGES, S. BOTTANI, G. BATT, P. HERSEN.*Long-term model predictive control of gene expression at the population and single-cell levels*, in "Proceedings of the National Academy of Sciences", 2012, vol. 109, n^{o} 35, p. 14271–14276

Project-Team SELECT

Model selection in statistical learning

IN COLLABORATION WITH: Laboratoire de mathématiques d'Orsay de l'Université de Paris-Sud (LMO)

IN PARTNERSHIP WITH:
CNRS

Université Paris-Sud (Paris 11)

RESEARCH CENTER
Saclay - Île-de-France

THEME
Optimization, machine learning and statistical methods

Table of contents

Project-Team SELECT

Creation of the Project-Team: 2007 January 01, end of the Project-Team: 2018 December 31

Keywords:

Computer Science and Digital Science:
A3.1.1. - Modeling, representation
A3.1.8. - Big data (production, storage, transfer)
A3.2.2. - Knowledge extraction, cleaning
A3.3.2. - Data mining
A3.3.3. - Big data analysis
A3.4.1. - Supervised learning
A3.4.2. - Unsupervised learning
A3.4.3. - Reinforcement learning
A3.4.4. - Optimization and learning
A3.4.5. - Bayesian methods
A3.4.6. - Neural networks
A3.4.7. - Kernel methods
A3.4.8. - Deep learning
A5.3.3. - Pattern recognition
A6.2.4. - Statistical methods
A6.2.6. - Optimization

Other Research Topics and Application Domains:
B1.1.4. - Genetics and genomics
B1.1.7. - Bioinformatics
B1.1.8. - Mathematical biology
B9.5.2. - Mathematics

1. Team, Visitors, External Collaborators

Research Scientists
Kevin Bleakley [Inria, Researcher]
Gilles Celeux [Inria, Emeritus]
Matthieu Lerasle [CNRS, Researcher]

Faculty Members
Pascal Massart [Team leader, Univ Paris-Sud, Professor]
Sylvain Arlot [Univ Paris-Sud, Professor]
Christine Keribin [Univ Paris-Sud, Associate Professor]
Patrick Pamphile [Univ Paris-Sud, Associate Professor]
Jean-Michel Poggi [Univ René Descartes, Professor, HDR]

Technical Staff
Benjamin Auder [CNRS]
Christian Poli [Inria]

PhD Students

Florence Ducros [Univ Paris-Sud, until Sep 2018]
Neska El Haouij [Univ Paris-Sud, until Sep 2018]
Benjamin Goehry [Univ Paris-Sud]
Hedi Hadiji [Univ Paris-Sud]
Minh Lien Nguyen [Univ Paris-Sud, until Oct 2018]

Administrative Assistant

Maeva Jeannot [Inria]

2. Overall Objectives

2.1. Model selection in Statistics

The research domain for the SELECT project is statistics. Statistical methodology has made great progress over the past few decades, with a variety of statistical learning software packages that support many different methods and algorithms. Users now face the problem of choosing among them, to select the most appropriate method for their data sets and objectives. The problem of model selection is an important but difficult problem, both theoretically and practically. Classical model selection criteria, which use penalized minimum-contrast criteria with fixed penalties, are often based on unrealistic assumptions.

SELECT aims to provide efficient model selection criteria with data-driven penalty terms. In this context, SELECT aims to improve the toolkit of statistical model selection criteria from both theoretical and practical perspectives. Currently, SELECT is focusing its effort on variable selection in statistical learning, hidden-structure models and supervised classification. Its domains of application concern reliability, curve classification, phylogenetic analysis and classification in genetics. New developments in SELECT activities are concerned with applications in biostatistics (statistical analysis of medical images) and biology.

3. Research Program

3.1. General presentation

From applications we treat on a day-to-day basis, we have learned that some assumptions currently used in asymptotic theory for model selection are often irrelevant in practice. For instance, it is not realistic to assume that the target belongs to the family of models in competition. Moreover, in many situations, it is useful to make the size of the model depend on the sample size, which makes asymptotic analyses breakdown. An important aim of SELECT is to propose model selection criteria which take such practical constraints into account.

3.2. A nonasymptotic view of model selection

An important goal of SELECT is to build and analyze penalized log-likelihood model selection criteria that are efficient when the number of models in competition grows to infinity with the number of observations. Concentration inequalities are a key tool for this, and lead to data-driven penalty choice strategies. A major research direction for SELECT consists of deepening the analysis of data-driven penalties, both from the theoretical and practical points of view. There is no universal way of calibrating penalties, but there are several different general ideas that we aim to develop, including heuristics derived from Gaussian theory, special strategies for variable selection, and resampling methods.

3.3. Taking into account the modeling purpose in model selection

Choosing a model is not only difficult theoretically. From a practical point of view, it is important to design model selection criteria that accommodate situations in which the data probability distribution P is unknown, and which take the model user's purpose into account. Most standard model selection criteria assume that P belongs to one of a set of models, without considering the purpose of the model. By also considering the model user's purpose, we can avoid or overcome certain theoretical difficulties, and produce flexible model selection criteria with data-driven penalties. The latter is useful in supervised classification and hidden-structure models.

3.4. Bayesian model selection

The Bayesian approach to statistical problems is fundamentally probabilistic: a joint probability distribution is used to describe the relationships among all unknowns and the data. Inference is then based on the posterior distribution, i.e., the conditional probability distribution of the parameters given the observed data. Exploiting the internal consistency of the probability framework, the posterior distribution extracts relevant information in the data and provides a complete and coherent summary of post-data uncertainty. Using the posterior to solve specific inference and decision problems is then straightforward, at least in principle.

4. Application Domains

4.1. Introduction

A key goal of SELECT is to produce methodological contributions in statistics. For this reason, the SELECT team works with applications that serve as an important source of interesting practical problems and require innovative methodology to address them. Many of our applications involve contracts with industrial partners, e.g., in reliability, although we also have several academic collaborations, e.g., in genetics and image analysis.

4.2. Curve classification

The field of classification for complex data such as curves, functions, spectra and time series, is an important problem in current research. Standard data analysis questions are being looked into anew, in order to define novel strategies that take the functional nature of such data into account. Functional data analysis addresses a variety of applied problems, including longitudinal studies, analysis of fMRI data, and spectral calibration.

We are focused in particular on unsupervised classification. In addition to standard questions such as the choice of the number of clusters, the norm for measuring the distance between two observations, and vectors for representing clusters, we must also address a major computational problem: the functional nature of the data, which requires new approaches.

4.3. Computer experiments and reliability

For several years now, SELECT has collaborated with the EDF-DER *Maintenance des Risques Industriels* group. One important theme involves the resolution of inverse problems using simulation tools to analyze incertainty in highly complex physical systems.

The other major theme concerns reliability, through a research collaboration with Nexter involving a Cifre convention. This collaboration concerns a lifetime analysis of a vehicle fleet to assess ageing.

Moreover, a collaboration is ongoing with Dassault Aviation on the modal analysis of mechanical structures, which aims to identify the vibration behavior of structures under dynamic excitation. From the algorithmic point of view, modal analysis amounts to estimation in parametric models on the basis of measured excitations and structural response data. In literature and existing implementations, the model selection problem associated with this estimation is currently treated by a rather weighty and heuristic procedure. In the context of our own research, model selection via penalization methods are being tested on this model selection problem.

4.4. Analysis of genomic data

For many years now, SELECT collaborates with Marie-Laure Martin-Magniette (URGV) for the analysis of genomic data. An important theme of this collaboration is using statistically sound model-based clustering methods to discover groups of co-expressed genes from microarray and high-throughput sequencing data. In particular, identifying biological entities that share similar profiles across several treatment conditions, such as co-expressed genes, may help identify groups of genes that are involved in the same biological processes.

SELECT collaborates with Anavaj Sakuntabhai and Philippe Dussart (Pasteur Institute) on predicting dengue severity using only low-dimensional clinical data obtained at hospital arrival. Further collaborations are underway in dengue fever and encephalitis with researchers at the Pasteur Institute, including with Jean-David Pommier.

SELECT is involved in the ANR "jeunes chercheurs" MixStatSeq directed by Cathy Maugis (INSA Toulouse), which is concerned with statistical analysis and clustering of RNASeq genomics data.

4.5. Pharmacovigilance

A collaboration is ongoing with Pascale Tubert-Bitter, Ismael Ahmed and Mohamed Sedki (Pharmacoepidemiology and Infectious Diseases, PhEMI) for the analysis of pharmacovigilance data. In this framework, the goal is to detect, as soon as possible, potential associations between certain drugs and adverse effects, which appeared after the authorized marketing of these drugs. Instead of working on aggregate data (contingency table) like is usually the case, the approach developed aims to deal with individual's data, which perhaps gives more information.

4.6. Spectroscopic imaging analysis of ancient materials

Ancient materials, encountered in archaeology and paleontology are often complex, heterogeneous and poorly characterized before physico-chemical analysis. A popular technique to gather as much physico-chemical information as possible, is spectro-microscopy or spectral imaging, where a full spectra, made of more than a thousand samples, is measured for each pixel. The produced data is tensorial with two or three spatial dimensions and one or more spectral dimensions, and requires the combination of an "image" approach with a "curve analysis" approach. Since 2010 SELECT, collaborates with Serge Cohen (IPANEMA) on clustering problems, taking spatial constraints into account.

5. New Software and Platforms

5.1. BlockCluster

Block Clustering
KEYWORDS: Statistic analysis - Clustering package
SCIENTIFIC DESCRIPTION: Simultaneous clustering of rows and columns, usually designated by biclustering, co-clustering or block clustering, is an important technique in two way data analysis. It consists of estimating a mixture model which takes into account the block clustering problem on both the individual and variables sets. The blockcluster package provides a bridge between the C++ core library and the R statistical computing environment. This package allows to co-cluster binary, contingency, continuous and categorical data-sets. It also provides utility functions to visualize the results. This package may be useful for various applications in fields of Data mining, Information retrieval, Biology, computer vision and many more.
FUNCTIONAL DESCRIPTION: BlockCluster is an R package for co-clustering of binary, contingency and continuous data based on mixture models.

- Participants: Christophe Biernacki, Gilles Celeux, Parmeet Bhatia, Serge Iovleff, Vincent Brault and Vincent Kubicki
- Partner: Université de Technologie de Compiègne
- Contact: Serge Iovleff
- URL: http://cran.r-project.org/web/packages/blockcluster/index.html

5.2. MASSICCC

Massive Clustering with Cloud Computing
KEYWORDS: Statistic analysis - Big data - Machine learning - Web Application

SCIENTIFIC DESCRIPTION: The web application let users use several software packages developed by Inria directly in a web browser. Mixmod is a classification library for continuous and categorical data. MixtComp allows for missing data and a larger choice of data types. BlockCluster is a library for co-clustering of data. When using the web application, the user can first upload a data set, then configure a job using one of the libraries mentioned and start the execution of the job on a cluster. The results are then displayed directly in the browser allowing for rapid understanding and interactive visualisation.
FUNCTIONAL DESCRIPTION: The MASSICCC web application offers a simple and dynamic interface for analysing heterogeneous data with a web browser. Various software packages for statistical analysis are available (Mixmod, MixtComp, BlockCluster) which allow for supervised and supervised classification of large data sets.

- Contact: Christophe Biernacki
- URL: https://massiccc.lille.inria.fr

5.3. Mixmod

Many-purpose software for data mining and statistical learning
KEYWORDS: Data mining - Classification - Mixed data - Data modeling - Big data
FUNCTIONAL DESCRIPTION: Mixmod is a free toolbox for data mining and statistical learning designed for large and highdimensional data sets. Mixmod provides reliable estimation algorithms and relevant model selection criteria.

It has been successfully applied to marketing, credit scoring, epidemiology, genomics and reliability among other domains. Its particularity is to propose a model-based approach leading to a lot of methods for classification and clustering.

Mixmod allows to assess the stability of the results with simple and thorough scores. It provides an easy-to-use graphical user interface (mixmodGUI) and functions for the R (Rmixmod) and Matlab (mixmodForMatlab) environments.

- Participants: Benjamin Auder, Christophe Biernacki, Florent Langrognet, Gérard Govaert, Gilles Celeux, Remi Lebret and Serge Iovleff
- Partners: CNRS - Université Lille 1 - LIFL - Laboratoire Paul Painlevé - HEUDIASYC - LMB
- Contact: Gilles Celeux
- URL: http://www.mixmod.org

6. New Results

6.1. Model selection in Regression and Classification

Participants: Gilles Celeux, Pascal Massart, Sylvain Arlot, Jean-Michel Poggi, Kevin Bleakley.

In collaboration with Damien Garreau, Sylvain Arlot studied the kernel change-point algorithm (KCP) proposed by Arlot, Celisse and Harchaoui (2012), which aims at locating an unknown number of change-points in the distribution of a sequence of independent data taking values in an arbitrary set. The change-points are selected by model selection with a penalized kernel empirical criterion. We provide a non-asymptotic result showing that, with high probability, the KCP procedure retrieves the correct number of change-points, provided that the constant in the penalty is well-chosen; in addition, KCP estimates the change-points location at the optimal rate. As a consequence, when using a characteristic kernel, KCP detects all kinds of change in the distribution (not only changes in the mean or the variance), and it is able to do so for complex structured data (not necessarily in Rd). Most of the analysis is conducted assuming that the kernel is bounded; part of the results can be extended when we only assume a finite second-order moment.

The well-documented and consistent variable selection procedure in model-based cluster analysis and classification that Cathy Maugis (INSA Toulouse) designed during her PhD thesis in SELECT, makes use of stepwise algorithms which are painfully slow in high dimensions. In order to circumvent this drawback, Gilles Celeux, in collaboration with Mohammed Sedki (Université Paris XI) and Cathy Maugis, have recently submitted an article where variables are sorted using a lasso-like penalization adapted to the Gaussian mixture model context. Using this ranking to select variables, they avoid the combinatory problem of stepwise procedures. The performances on challenging simulated and real data sets are similar to the standard procedure, with a CPU time divided by a factor of more than a hundred.

In collaboration with Benjamin Charlier and Jean-Michel Marin (Université de Montpellier), Gilles Celeux has started research aiming to propose a rapid Bayesian algorithm to estimate simply the mode of a posterior distribution for hidden structure models. This Bayesian procedure is of interest for two reasons. First, it leads to regularised estimation, which is useful for poorly posed problems. Second, it is an interesting alternative to variational approximation.

6.2. Estimator selection and statistical tests

Participant: Sylvain Arlot.

Sylvain Arlot wrote a book chapter about cross-validation in 2018. This text defines all classical cross-validation procedures, and studies their properties for two different goals: estimating the risk of a given estimator, and selecting the best estimator among a given family. For the risk estimation problem, it computes the bias (which can also be corrected) and the variance of cross-validation methods. For estimator selection, it first provides a first-order analysis (based on expectations). Then, it explains how to take into account second-order terms (from variance computations, and by taking into account the usefulness of over-penalization). This allows, in the end, to provide some guidelines for choosing the best cross-validation method for a given learning problem.

6.3. Statistical learning methodology and theory

Participants: Gilles Celeux, Serge Cohen, Christine Keribin, Michel Prenat, Sylvain Arlot, Benjamin Auder, Jean-Michel Poggi, Neska El Haouij, Kevin Bleakley, Matthieu Lerasle.

Sylvain Arlot wrote a book chapter about supervised statistical learning, from the mathematical point of view in 2018. This text describes the general prediction problem and the two key examples of regression and binary classification. Then, it studies two kinds of learning rules: empirical risk minimizers, which naturally lead to convex risks in classification, and local averaging rules, for which a universal consistency result can be obtained. Finally, it identifies the limits of learning in order to underline its challenges. The text ends with some useful probabilistic tools and some exercises.

Gilles Celeux and Serge Cohen have started research in collaboration with Agnès Grimaud (UVSQ) to perform clustering of hyperspectral images which respects spatial constraints. This is a one-class classification problem where distances between spectral images are given by the χ^2 distance, while spatial homogeneity is associated with a single link distance. This year they have developed a hybrid hierarchical clustering procedure in which sub-clusters respecting spatial consistency are constructed. Then, these sub-clusters are merged without taking spatial constraints into account. This strategy leads to a more realistic segmentation of spectral images.

Gilles Celeux continued his collaboration with Jean-Patrick Baudry on model-based clustering. Last year, they started work on assessing model-based clustering methods on cytometry data sets. The interest of these is that they involve combining clustering and classification tasks in a unified framework. This year, this work was completed, and performed well in comparison with state-of-the-art procedures.

Gillies Celeux has continued research on missing data for model-based clustering in collaboration with Christophe Biernacki (Modal team, Inria Lille) and Julie Josse (École Polytechnique). This year, they implemented several algorithms to estimate their logistic model for mixture analysis involving not missing-at-random mixtures.

In the framework of MASSICCC, Benjamin Auder and Gilles Celeux have started research on the graphical representation of model-based clusters. The aim of this is to better-display proximity between clusters.It leads to a simple procedure to represent the proximity between clusters without any additional assumptions.

After having proved the consistency and asymptotic normality of Latent Block Model estimators with V. Brault and M. Mariadassou, Christine Keribin has worked on the behavior of the ICL and BIC model criteria in this model, and in particular on their probable asymptotic equivalence.

Christine Keribin has started a new collaboration with Christophe Biernacki (Inria Modal Team) to study the ability for co-clustering to be a good regularized method for clustering in HD, which was presented at the CMStatistics 2018 conference.

J-M. Poggi (with R. Genuer), published a survey paper dedicated to “Arbres CART et Forêts aléatoires, Importance et sélection de variables”, as a book chapter published in: “Apprentissage Statistique et Données Massives” by Technip.

J.-M. Poggi and N. El Haouij (with R. Ghozi, S. Sevestre Ghalila and M. Jaïdane) provide a random forest-based method for the selection of physiological functional variables in order to classify the stress level during real-world driving experience. The contribution of this study is twofold: on the methodological side, it considers physiological signals as functional variables and offers a procedure of data processing and variable selection. On the applied side, the proposed method provides a “blind” procedure of driver’s stress level classification that does not depend on the expert-based studies of physiological signals. This work has been published in Statistical Methods & Applications.

J.-M. Poggi and N. El Haouij (with R. Ghozi, S. Sevestre Ghalila and M. Jaïdane provide a system and database to assess driver’s attention, called aAffectiveROAD. A paper presenting it has been published in the proceedings of the 33rd ACM Symposium on Applied Computing SAC’18.

6.4. Statistical analysis of genomic data

Participant: Kevin Bleakley.

In collaboration with Benno Schwikowski, Iryna Nikolayeva and Anavaj Sakuntabhai (Pasteur Institute, Paris), Kevin Bleakley worked on using 2-d isotonic regression to predict dengue fever severity at hospital arrival using high-dimensional microarray gene expression data. Important marker genes for dengue severity have been detected, some of which now have been validated in external lab trials, and an article on this was published in the Journal of Infectious Diseases in 2018.

Kevin Bleakley has also collaborated with Inserm/Paris-Saclay researchers at Kremlin-Bicêtre hospital on cyclic transcriptional clocks and renal corticosteroid signaling, and has developed novel statistical tests for detecting synchronous signals. This work was published in the FASEB journal in 2018.

Kevin Bleakley worked as part of a consortium on a crowdsourced Dream Challenge in 2018 on using molecular signatures to predict susceptibility to viral infection. Essentially, many teams of researchers from around the world used machine learning (statistical learning) algorithms to learn on training data then test on unseen real data. In the final stage, methods from several teams were combined to improve overall prediction performance. The article “A crowdsourced analysis to identify ab initio molecular signatures predictive of susceptibility to viral infection” was published in Nature Communications in 2018.

6.5. Reliability

Participants: Gilles Celeux, Florence Ducros, Patrick Pamphile.

From June 2015 until June 2018 when she defended it, in the framework of a CIFRE convention with Nexter, Florence Ducros researched a thesis on the modeling of aging of vehicles, supervised by Gilles Celeux and Patrick Pamphile. This thesis should lead to designing an efficient maintenance strategy according to vehicle use profiles. Moreover, warranty cost calculations are made in the context of heterogeneous usages. This required estimations of mixtures and competing risk models in a highly-censored setting.

This year, Patrick Pamphile and Florence Ducros have published an article which proposes a two-component Weibull mixture model for modelling unobserved heterogeneity in heavily censored lifetime data collection. Performance of classical estimation methods (maximum of likelihood, EM, full Bayes and MCMC) are poor due to the high number of parameters and the heavy censoring. Thus, a Bayesian bootstrap method called Bayesian Restoration Maximization, was used. Sampling from the posterior distribution was obtained thanks to an importance sampling technique. Simulation results showed that, even with heavy censoring, BRM is effective both in term of estimate's precision and computation times.

6.6. Dynamical systems

Participant: Sylvain Arlot.

In collaboration with Stefano Marmi and Duccio Papini, Sylvain Arlot proposed a new model for the time evolution of livestock commodities which exhibits endogenous deterministic stochastic behaviour. The model is based on the Yoccoz-Birkeland integral equation, a model first developed for studying the time-evolution of single species with high average fertility, a relatively short mating season and density dependent reproduction rates. This equation is then coupled with a differential equation describing the price of a livestock commodity driven by the unbalance between its demand and supply. At its birth the cattle population is split into two parts: reproducing females and cattle for butchery. The relative amount of the two is determined by the spot price of the meat. We prove the existence of an attractor and we investigate numerically its properties: the strange attractor existing for the original Yoccoz-Birkeland model is persistent but its chaotic behaviour depends also from the price evolution in an essential way.

6.7. Soccer forecasting

Participants: Gilles Celeux, Jean-Louis Foulley.

In collaboration with Jean-Louis Foulley (Montpellier University), Gilles Celeux has proposed a penalty criterion for assessing correct score forecasting in soccer matches. They have defined the subject of a Masters internship for next year to predict scores of soccer matches via Poisson models using maximum likelihood and Bayesian inference.

6.8. Electricity load forecasting and clustering

Participants: Jean-Michel Poggi, Benjamin Auder, Benjamin Goehry.

B. Auder, J-M. Poggi (with J. Cugliari, Y. Goude) are interested in hierarchical time-series for bottom-up forecasting. The idea is to disaggregate the signal in such a way that the sum of disaggregated forecasts improves the direct prediction. The 3-steps strategy defines numerous super-consumers by curve clustering, builds a hierarchy of partitions and selects the best one minimizing a forecast criterion. Using a nonparametric model to handle forecasting, and wavelets to define various notions of similarity between load curves, this disaggregation strategy applied to French individual consumers leads to a gain of 16% in forecast accuracy. Then the upscaling capacity of this strategy facing massive data is explored and different proposals using R are experimented. The proposed solutions to make the algorithm scalable combines data storage, parallel computing and double clustering step to define the super-consumers. This has been published in the journal Energies.

Benjamin Goehry is completing a thesis co-supervised by P. Massart and J-M. Poggi, aiming at extending this scheme by introducing the use of random forests as time series forecasting models adapted to each cluster.

J.-M. Poggi (with J. Cugliari) published in Wiley StatsRef-Statistics Reference Online, a paper entitled Electricity demand forecasting. the focus is on short-term demand forecasting at some aggregate level (e.g., zone or nationwide demands) from data with at least hourly sampled data. The main salient features of the load curve are first highlighted. Some of the common covariates used in the prediction task are also discussed. Then, some basic or now classical methodological approaches for electricity demand forecasting are detailed.

7. Bilateral Contracts and Grants with Industry

7.1. Contract with NEXTER

Participants: Gilles Celeux, Florence Ducros, Patrick Pamphile.

SELECT has a contract with Nexter regarding modeling the reliability of vehicles.

8. Partnerships and Cooperations

8.1. Regional Initiatives

Sylvain Arlot and Pascal Massart co-organize a working group at ENS (Ulm) on statistical learning.

8.2. National Initiatives

8.2.1. ANR

SELECT is part of the ANR-funded MixStatSeq.

Sylvain Arlot and Matthieu Lerasle are part of the ANR grant FAST-BIG (Efficient Statistical Testing for high-dimensional Models: application to Brain Imaging and Genetics), which is lead by Bertrand Thirion (Inria Saclay, Parietal).

8.3. International Initiatives

Gilles Celeux is one of the co-organizers of the international working group on model-based clustering. This year this workshop took place in Ann Arbor (USA).

8.4. International Research Visitors

8.4.1. Visits to International Teams

8.4.1.1. Research Stays Abroad

Kevin Bleakley stayed at the Pasteur Institute, Cambodia, while working on several collaborations in dengue fever and encephalitis, from February–March 2018.

Jean-Michel Poggi: Universidad de la República (Montevideo, Uruguay), Facultad de Ingeniería, Instituto de Matemática y Estadística “Prof. Ing. Rafael Laguardia”, 17-28 February 2018.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

Sylvain Arlot co-organized (with Christophe Giraud and Gilles Stoltz) the conference “Deux complices en statistique” (Two days in honor of Pascal Massart and Lucien Birgé), at IHES (Bures-sur-Yvette) and IMO (Orsay).

9.1.1.2. Member of the Organizing Committees

- Gilles Celeux is one of the co-organizers of the international working group on model-based clustering. This year the workshop took place in Ann Arbor, USA.
- Sylvain Arlot is one of the co-organizers of the Junior Conference on Data Science and Engineering at Paris-Saclay (3nd edition in 2018).

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

Jean-Michel Poggi was a member of the Scientific Programme Committee, ENBIS 2018, Nancy, 2-6, September 2018, and member of the Scientific Committee of the summer school on Clustering, Data, Analysis And Visualization Of Complex Data, May 21-25, 2018.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

Gilles Celeux is Editor-in-Chief of the *Journal de la SFdS*. He is Associate Editor of *Statistics and Computing*, *CSBIGS*.

Pascal Massart is Associate Editor of *Annals of Statistics*, *Confluentes Mathematici*, and *Foundations and Trends in Machine Learning*.

Jean-Michel Poggi is Associate Editor of *Journal of Statistical Software*, *Journal de la SFdS*, and *CSBIGS*.

Sylvain Arlot is associate editor for the Annales de l'Institut Henri Poincaré B, Probability and Statistics.

9.1.3.2. Reviewer - Reviewing Activities

The members of the team have reviewed numerous papers for numerous international journals.

9.1.4. Invited Talks

The members of the team have given many invited talks on their research in the course of 2018.

9.1.5. Leadership within the Scientific Community

Jean-Michel Poggi is:

- Vice-President of ECAS (European Courses in Advanced Statistics)
- Council Member of the ISI (2015-19)
- Member of the Board of Directors of the ERS of IASC (since 2014)
- Council member of FENStatS (Federation of European National Statistical Societies)

9.1.6. Scientific Expertise

Jean-Michel Poggi was member of the Box Medal committee for 2018, and in the jury for the Marie-Jeanne Laurent Duhamel prize (SFdS).

9.1.7. Research Administration

Jean-Michel Poggi is the vice-president of ECAS (European Courses in Advanced Statistics) since 2015.

Sylvain Arlot coordinates (jointly with Marc Schoenauer, Inria Saclay) the math-STIC program of the Labex Mathématique Hadamard.

Christine Keribin is treasurer of the Société Française de Statistique (SFdS).

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

SELECT members teach various courses at several different universities, and in particular the Master 2 “Mathématique de l'aléatoire” of Université Paris-Saclay.

9.2.2. Supervision

PhD: Neska El Haouij, 2014, Jean-Michel Poggi, Meriem Jaïdane, Raja Ghozi (ENIT Tunisie), and Sylvie Sevestre-Ghalila (CEA LinkLab). Defended in July 2018.

PhD: Florence Ducros, 2015, Gilles Celeux and Patrick Pamphile. Defended June 2018.

PhD in progress: Claire Brécheteau, 2015, Pascal Massart

PhD in progress: Hedi Hadiji, 2017, Pascal Massart

PhD in progress: Guillaume Maillard, 2016, Sylvain Arlot and Matthieu Lerasle

PhD in progress: Jeanne Nguyen, 2015, Claire Lacour and Vincent Rivoirard (Univ Paris Dauphine)

PhD in progress: Benjamin Goehry, 2015, Pascal Massart and Jean-Michel Poggi

PhD in progress: Tuan-Binh Nguyen, 2018, Sylvain Arlot and Bertrand Thirion

9.2.3. Juries

HDR: Emilie Lebarbier (Pascal Massart, referee; Sylvain Arlot, president)

Members of SELECT have participated in numerous juries during 2018.

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] G. CELEUX, C. MAUGIS-RABUSSEAU, M. SEDKI.*Variable selection in model-based clustering and discriminant analysis with a regularization approach*, in "Advances in Data Analysis and Classification", 2018 [*DOI :* 10.1007/S11634-018-0322-5], https://hal.inria.fr/hal-01053784

[2] F. DUCROS, P. PAMPHILE.*Bayesian estimation of Weibull mixture in heavily censored data setting*, in "Reliability Engineering and System Safety", December 2018, https://hal.archives-ouvertes.fr/hal-01645618

[3] S. FOURATI, A. TALLA, M. MAHMOUDIAN, J. BURKHART, R. KLÉN, R. HENAO, T. YU, Z. AYDIN, K. Y. YEUNG, M. E. AHSEN, R. ALMUGBEL, S. JAHANDIDEH, X. LIANG, T. NORDLING, M. SHIGA, A. STANESCU, R. VOGEL, G. PANDEY, C. CHIU, M. MCCLAIN, C. WOODS, G. GINSBURG, L. ELO, E. TSALIK, L. MANGRAVITE, S. SIEBERTS, K. BLEAKLEY.*A crowdsourced analysis to identify ab initio molecular signatures predictive of susceptibility to viral infection*, in "Nature Communications", December 2018, vol. 9, n^{o} 1, https://hal.inria.fr/hal-01960254

[4] D. GARREAU, S. ARLOT.*Consistent change-point detection with kernels*, in "Electronic journal of statistics ", December 2018, vol. 12, n^{o} 2, p. 4440-4486, https://arxiv.org/abs/1612.04740 , https://hal.archives-ouvertes.fr/hal-01416704

[5] F. LE BILLAN, L. AMAZIT, K. BLEAKLEY, Q.-Y. XUE, E. PUSSARD, C. LHADJ, P. KOLKHOF, S. VIENGCHAREUN, J. FAGART, M. LOMBES.*Corticosteroid receptors adopt distinct cyclical transcriptional signatures*, in "FASEB Journal", October 2018, vol. 32, n^{o} 10, p. 5626-5639, https://hal.inria.fr/hal-01960262

[6] I. NIKOLAYEVA, P. BOST, I. CASADEMONT, V. DUONG, F. KOETH, M. PROT, U. CZERWINSKA, S. LY, K. BLEAKLEY, T. CANTAERT, P. DUSSART, P. BUCHY, E. SIMON-LORIERE, A. SAKUNTABHAI, B. SCHWIKOWSKI.*A Blood RNA Signature Detecting Severe Disease in Young Dengue Patients at Hospital Arrival*, in "Journal of Infectious Diseases", June 2018, vol. 217, n^{o} 11, p. 1690-1698 [*DOI :* 10.1093/INFDIS/JIY086], https://hal.inria.fr/hal-01960247

[7] L. Raynal, J.-M. Marin, P. Pudlo, M. Ribatet, C. Robert, A. Estoup, O. Stegle.*ABC random forests for Bayesian parameter inference*, in "Bioinformatics", October 2018, https://hal.archives-ouvertes.fr/hal-01961126

Invited Conferences

[8] C. Keribin, G. Celeux, V. Robert.*The Latent Block Model: a useful model for high dimensional data*, in "Mixture models : Theory and applications", Paris, France, June 2018, https://hal.archives-ouvertes.fr/hal-01957710

International Conferences with Proceedings

[9] F. Chazal, V. Divol.*The density of expected persistence diagrams and its kernel based estimation*, in "Symposium of Computational Geometry (SoCG 2018)", Budapest, Hungary, June 2018, Extended version of a paper to appear in the proceedings of the Symposium of Computational Geometry 2018, https://hal.archives-ouvertes.fr/hal-01716181

Conferences without Proceedings

[10] C. Biernacki, B. Auder, G. Celeux, J. Demont, F. Langrognet, V. Kubicki, C. Poli, J. Renault.*MASSICCC: A SaaS Platform for Clustering and Co-Clustering of Mixed Data*, in "Workshop MixStatSeq: “Mixture models: Theory and applications”", Paris, France, June 2018, https://hal.archives-ouvertes.fr/hal-01949175

[11] C. Biernacki, G. Celeux, J. Josse, F. Laporte.*Dealing with missing data in model-based clustering through a MNAR model*, in "CMStatistics 2018 - 11th International Conference of the ERCIM WG on Computational and Methodological Statistics", Pise, Italy, December 2018, https://hal.archives-ouvertes.fr/hal-01949120

[12] C. Keribin, C. Biernacki.*Co-clustering: A versatile way to perform clustering in high dimension*, in "The 11th International Conference of the ERCIM WG on Computational and Methodological Statistics (CMStatistics 2018)", Pise, Italy, December 2018, https://hal.archives-ouvertes.fr/hal-01949116

Scientific Books (or Scientific Book chapters)

[13] S. Arlot.*Cross-validation*, in "Apprentissage statistique et données massives", M. Maumy-Bertrand, G. Saporta, C. Thomas-Agnan (editors), Editions Technip, May 2018, https://arxiv.org/abs/1703.03167 , https://hal.archives-ouvertes.fr/hal-01485508

[14] S. Arlot.*Tutorial on statistical learning*, in "Apprentissage statistique et données massives", M. Maumy-Bertrand, G. Saporta, C. Thomas-Agnan (editors), Editions Technip, May 2018, https://hal.archives-ouvertes.fr/hal-01485506

[15] G. Celeux, S. Frühwirth-Schnatter, C. Robert.*Model Selection for Mixture Models-Perspectives and Strategies*, in "Handbook of Mixture Analysis", CRC Press, December 2018, https://hal.archives-ouvertes.fr/hal-01961077

[16] S. Frühwirth-Schnatter, G. Celeux, C. P. Robert.*Handbook of Mixture Analysis*, Chapman & Hall/CRC Handbooks of Modern Statistical Methods, CRC Press, November 2018, p. 1-536, https://hal.inria.fr/hal-01928103

[17] C. ROBERT, G. CELEUX, K. KAMARY, G. MALSINER-WALLI, J.-M. MARIN.*Computational Solutions for Bayesian Inference in Mixture Models*, in "Handbook of Mixture Analysis", CRC Press, December 2018, https://arxiv.org/abs/1812.07240 , https://hal.archives-ouvertes.fr/hal-01961038

Research Reports

[18] Y. AUFFRAY.*A variational interpretation of classification EM*, Dassault Aviation, September 2018, https://hal.inria.fr/hal-01882980

[19] M. LERASLE, Z. SZABÓ, T. MATHIEU, G. LECUÉ.*MONK – Outlier-Robust Mean Embedding Estimation by Median-of-Means*, CNRS / LMO - Laboratoire de Mathématiques d'Orsay, Orsay ; Ecole Polytechnique (Palaiseau, France) ; Laboratoire de Mathématiques d'Orsay ; ENSAE ParisTech ; Inria Saclay, équipe SELECT, October 2018, https://hal.archives-ouvertes.fr/hal-01705881

Other Publications

[20] E. AAMARI, J. KIM, F. CHAZAL, B. MICHEL, A. RINALDO, L. WASSERMAN.*Estimating the Reach of a Manifold*, January 2018, https://arxiv.org/abs/1705.04565 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01521955

[21] S. ARLOT, S. MARMI, D. PAPINI.*Coupling the Yoccoz-Birkeland population model with price dynamics: chaotic livestock commodities market cycles*, May 2018, https://arxiv.org/abs/1803.05404 - 26 pages, 19 figures, https://hal.archives-ouvertes.fr/hal-01812794

[22] M. BAELDE, C. BIERNACKI, R. GREFF.*Real-Time Monophonic and Polyphonic Audio Classification from Power Spectra*, July 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01834221

[23] C. BRÉCHETEAU, A. FISCHER, C. LEVRARD.*Robust Bregman Clustering*, December 2018, https://arxiv.org/abs/1812.04356 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01948051

[24] G. CHINOT, G. LECUÉ, M. LERASLE.*Statistical learning with Lipschitz and convex loss functions*, November 2018, https://arxiv.org/abs/1810.01090 - working paper or preprint [*DOI :* 10.01090], https://hal.archives-ouvertes.fr/hal-01923033

[25] A. HAVET, M. LERASLE, É. MOULINES.*Density estimation for RWRE*, June 2018, https://arxiv.org/abs/1806.05839 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01815990

[26] S. LE CORFF, M. LERASLE, E. VERNET.*A Bayesian nonparametric approach for generalized Bradley-Terry models in random environment*, August 2018, https://arxiv.org/abs/1808.08104 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01860352

[27] G. LECUÉ, M. LERASLE.*Robust machine learning by median-of-means : theory and practice*, November 2018, https://arxiv.org/abs/1711.10306 - 48 pages, 6 figures, https://hal.archives-ouvertes.fr/hal-01923036

[28] G. LECUÉ, M. LERASLE, T. MATHIEU.*Robust classification via MOM minimization*, November 2018, https://arxiv.org/abs/1808.03106 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01923035

[29] Y. LIU, C. KERIBIN, T. POPOVA, Y. ROZENHOLC.*Statistical estimation of genomic alterations of tumors*, January 2018, working paper or preprint, https://hal.inria.fr/hal-01688885

[30] M.-L. J. NGUYEN.*Nonparametric method for sparse conditional density estimation in moderately large dimensions*, January 2018, https://arxiv.org/abs/1801.06477 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01688664

[31] C. ROBERT, G. CELEUX, J. JEWSON, J. JOSSE, J.-M. MARIN, C. P. ROBERT.*Some discussions on the Read Paper "Beyond subjective and objective in statistics" by A. Gelman and C. Hennig*, January 2019, https://arxiv.org/abs/1705.03727 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01968779

Project-Team SPECFUN

Symbolic Special Functions : Fast and Certified

RESEARCH CENTER
Saclay - Île-de-France

THEME
Algorithmics, Computer Algebra and Cryptology

Table of contents

Project-Team SPECFUN

Creation of the Team: 2012 November 01, updated into Project-Team: 2014 July 01

Keywords:

Computer Science and Digital Science:

A2.1.10. - Domain-specific languages
A2.1.11. - Proof languages
A2.4.3. - Proofs
A4.5. - Formal methods for security
A7.2. - Logic in Computer Science
A8.1. - Discrete mathematics, combinatorics
A8.3. - Geometry, Topology
A8.4. - Computer Algebra
A8.5. - Number theory

Other Research Topics and Application Domains:

B9.5.2. - Mathematics
B9.5.3. - Physics

1. Team, Visitors, External Collaborators

Research Scientists

Frédéric Chyzak [Team leader, Inria, Researcher, HDR]
Alin Bostan [Inria, Researcher, HDR]
Georges Gonthier [Inria, Senior Researcher]
Pierre Lairez [Inria, Researcher]

External Collaborators

Philippe Dumas [Ministère de l'Éducation Nationale (retired)]
Guy Fayolle [Inria, Senior Researcher (emeritus)]
Marc Mezzarobba [CNRS, Researcher]

Post-Doctoral Fellow

Svyatoslav Covanov [Inria, from Oct 2018]

Administrative Assistant

Stéphanie Aubin [Inria, until Sep 2018]

2. Overall Objectives

2.1. Scientific challenges, expected impact

The general orientation of our team is described by the short name given to it: *Special Functions*, that is, particular mathematical functions that have established names due to their importance in mathematical analysis, physics, and other application domains. Indeed, we ambition to study special functions with the computer, by combined means of computer algebra and formal methods.

Computer-algebra systems have been advertised for decades as software for "doing mathematics by computer" [60]. For instance, computer-algebra libraries can uniformly generate a corpus of mathematical properties about special functions, so as to display them on an interactive website. This possibility was recently shown by the computer-algebra component of the team [16]. Such an automated generation significantly increases the reliability of the mathematical corpus, in comparison to the content of existing static authoritative handbooks. The importance of the validity of these contents can be measured by the very wide audience that such handbooks have had, to the point that a book like [11] remains one of the most cited mathematical publications ever and has motivated the 10-year-long project of writing its successor [13]. However, can the mathematics produced "by computer" be considered as *true* mathematics? More specifically, whereas it is nowadays well established that the computer helps in discovering and observing new mathematical phenomenons, can the mathematical statements produced with the aid of the computer and the mathematical results computed by it be accepted as valid mathematics, that is, as having the status of mathematical *proofs*? Beyond the reported weaknesses or controversial design choices of mainstream computer-algebra systems, the issue is more of an epistemological nature. It will not find its solution even in the advent of the ultimate computer-algebra system: the social process of peer-reviewing just falls short of evaluating the results produced by computers, as reported by Th. Hales [39] after the publication of his proof of the Kepler Conjecture about sphere packing.

A natural answer to this deadlock is to move to an alternative kind of mathematical software and to use a proof assistant to check the correctness of the desired properties or formulas. The success of large-scale formalization projects, like the Four-Color Theorem of graph theory [34], the above-mentioned Kepler Conjecture [39], and the Odd Order Theorem of group theory [0], have increased the understanding of the appropriate software-engineering methods for this peculiar kind of programming. For computer algebra, this legitimates a move to proof assistants now.

The Dynamic Dictionary of Mathematical Functions [0] (DDMF) [16] is an online computer-generated handbook of mathematical functions that ambitions to serve as a reference for a broad range of applications. This software was developed by the computer-algebra component of the team as a project [0] of the MSR–INRIA Joint Centre. It bases on a library for the computer-algebra system Maple, Algolib [0], whose development started 20 years ago in ÃPI Algorithms [0]. As suggested by the constant questioning of certainty by new potential users, DDMF deserves a formal guarantee of correctness of its content, on a level that proof assistants can provide. Fortunately, the maturity of special-functions algorithms in Algolib makes DDMF a stepping stone for such a formalization: it provides a well-understood and unified algorithmic treatment, without which a formal certification would simply be unreachable.

The formal-proofs component of the team emanates from another project of the MSR–INRIA Joint Centre, namely the Mathematical Components project (MathComp) [0]. Since 2006, the MathComp group has endeavoured to develop computer-checked libraries of formalized mathematics, using the Coq proof assistant [56]. The methodological aim of the project was to understand the design methods leading to successful large-scale formalizations. The work culminated in 2012 with the completion of a formal proof of the Odd Order Theorem, resulting in the largest corpus of algebraic theories ever machine-checked with a proof assistant and a whole methodology to effectively combine these components in order to tackle complex formalizations. In particular, these libraries provide a good number of the many algebraic objects needed to reason about special functions and their properties, like rational numbers, iterated sums, polynomials, and a rich hierarchy of algebraic structures.

The present team takes benefit from these recent advances to explore the formal certification of the results collected in DDMF. The aim of this project is to concentrate the formalization effort on this delimited area, building on DDMF and the Algolib library, as well as on the Coq system [56] and on the libraries developed by the MathComp project.

[0] http://www.msr-inria.inria.fr/news/the-formalization-of-the-odd-order-theorem-has-been-completed-the-20-septembre-2012/
[0] http://ddmf.msr-inria.inria.fr/1.9.1/ddmf
[0] http://www.msr-inria.inria.fr/projects/dynamic-dictionary-of-mathematical-functions/
[0] http://algo.inria.fr/libraries/
[0] http://algo.inria.fr/
[0] http://www.msr-inria.fr/projects/mathematical-components/

2.1.1. Use computer algebra but convince users beyond reasonable doubt

The following few opinions on computer algebra are, we believe, typical of computer-algebra users' doubts and difficulties when using computer-algebra systems:

- Fredrik Johansson, expert in the multi-precision numerical evaluation of special functions and in fast computer-algebra algorithms, writes on his blog [45]: Mathematica is great for cross-checking numerical values, but it's not unusual to run into bugs, so *triple checking is a good habit*. One answer in the discussion is: We can claim that Mathematica has [...] *an impossible to understand semantics*: If Mathematica's output is wrong then change the input. If you don't like the answer, change the question. That seems to be the philosophy behind.
- A professor's advice to students [52] on using Maple: You may wish to use Maple to check your homework answers. If you do then keep in mind that Maple sometimes gives the *wrong answer, usually because you asked incorrectly, or because of niceties of analytic continuation*. You may even be bitten by an occasional Maple bug, though that has become fairly unlikely. Even with as powerful a tool as Maple you will still *have to devise your own checks* and you will still have to think.
- Jacques Carette, former head of the maths group at Maplesoft, about a bug [12] when asking Maple to take the limit `limit(f(n) * exp(-n), n = infinity)` for an undetermined function `f`: The problem is that there is an *implicit assumption in the implementation* that unknown functions do not 'grow too fast'.

As explained by the expert views above, complaints by computer-algebra users are often due to their misunderstanding of what a computer-algebra systems is, namely a purely syntactic tool for calculations, that the user must complement with a semantics. Still, robustness and consistency of computer-algebra systems are not ensured as of today, and, whatever Zeilberger may provocatively say in his Opinion 94 [61], a firmer logical foundation is necessary. Indeed, the fact is that many bugs in a computer-algebra system cannot be fixed by just the usual debugging method of tracking down the faulty lines in the code. It is sort of by design: assumptions that too often remain implicit are really needed by the design of symbolic algorithms and cannot easily be expressed in the programming languages used in computer algebra. A similar certification initiative has already been undertaken in the domain of numerical computing, in a successful manner [43], [19]. It is natural to undertake a similar approach for computer algebra.

2.1.2. Make computer algebra and formal proofs help one another

Some of the mathematical objects that interest our team are still totally untouched by formalization. When implementing them and their theory inside a proof assistant, we have to deal with the pervasive discrepancy between the published literature and the actual implementation of computer-algebra algorithms. Interestingly, this forces us to clarify our computer-algebraic view on them, and possibly make us discover holes lurking in published (human) proofs. We are therefore convinced that the close interaction of researchers from both fields, which is what we strive to maintain in this team, is a strong asset.

For a concrete example, the core of Zeilberger's creative telescoping manipulates rational functions up to simplifications. In summation applications, checking that these simplifications do not hide problematic divisions by 0 is most often left to the reader. In the same vein, in the case of integrals, the published algorithms do not check the convergence of all integrals, especially in intermediate calculations. Such checks are again left to the readers. In general, we expect to revisit the existing algorithms to ensure that they are meaningful for genuine mathematical sequences or functions, and not only for algebraic idealizations.

Another big challenge in this project originates in the scientific difference between computer algebra and formal proofs. Computer algebra seeks speed of calculation on *concrete instances* of algebraic data structures (polynomials, matrices, etc). For their part, formal proofs manipulate symbolic expressions in terms of *abstract variables* understood to represent generic elements of algebraic data structures. In view of this, a continuous challenge is to develop the right, hybrid thinking attitude that is able to effectively manage concrete and abstract values simultaneously, alternatively computing and proving with them.

2.1.3. Experimental mathematics with special functions

Applications in combinatorics and mathematical physics frequently involve equations of so high orders and so large sizes, that computing or even storing all their coefficients is impossible on existing computers. Making this tractable is an extraordinary challenge. The approach we believe in is to design algorithms of good—ideally quasi-optimal—complexity in order to extract precisely the required data from the equations, while avoiding the computationally intractable task of completely expanding them into an explicit representation.

Typical applications with expected high impact are the automatic discovery and algorithmic proof of results in combinatorics and mathematical physics for which human proofs are currently unattainable.

2.2. Research axes

The implementation of certified symbolic computations on special functions in the Coq proof assistant requires both investigating new formalization techniques and renewing the traditional computer-algebra viewpoint on these standard objects. Large mathematical objects typical of computer algebra occur during formalization, which also requires us to improve the efficiency and ergonomics of Coq. In order to feed this interdisciplinary activity with new motivating problems, we additionally pursue a research activity oriented towards experimental mathematics in application domains that involve special functions. We expect these applications to pose new algorithmic challenges to computer algebra, which in turn will deserve a formal-certification effort. Finally, DDMF is the motivation and the showcase of our progress on the certification of these computations. While striving to provide a formal guarantee of the correctness of the information it displays, we remain keen on enriching its mathematical content by developing new computer-algebra algorithms.

2.2.1. Computer algebra certified by the Coq system

Our formalization effort consists in organizing a cooperation between a computer-algebra system and a proof assistant. The computer-algebra system is used to produce efficiently algebraic data, which are later processed by the proof assistant. The success of this cooperation relies on the design of appropriate libraries of formalized mathematics, including certified implementations of certain computer-algebra algorithms. On the other side, we expect that scrutinizing the implementation and the output of computer-algebra algorithms will shed a new light on their semantics and on their correctness proofs, and help clarifying their documentation.

2.2.1.1. Libraries of formalized mathematics

The appropriate framework for the study of efficient algorithms for special functions is *algebraic*. Representing algebraic theories as Coq formal libraries takes benefit from the methodology emerging from the success of ambitious projects like the formal proof of a major classification result in finite-group theory (the Odd Order Theorem) [32].

Yet, a number of the objects we need to formalize in the present context has never been investigated using any interactive proof assistant, despite being considered as commonplaces in computer algebra. For instance there is up to our knowledge no available formalization of the theory of non-commutative rings, of the algorithmic theory of special-functions closures, or of the asymptotic study of special functions. We expect our future formal libraries to prove broadly reusable in later formalizations of seemingly unrelated theories.

2.2.1.2. Manipulation of large algebraic data in a proof assistant

Another peculiarity of the mathematical objects we are going to manipulate with the Coq system is their size. In order to provide a formal guarantee on the data displayed by DDMF, two related axes of research have to be pursued. First, efficient algorithms dealing with these large objects have to be programmed and run in Coq. Recent evolutions of the Coq system to improve the efficiency of its internal computations [14], [17] make this objective reachable. Still, how to combine the aforementioned formalization methodology with these cutting-edge evolutions of Coq remains one of the prospective aspects of our project. A second need is to help users *interactively* manipulate large expressions occurring in their conjectures, an objective for which little has been done so far. To address this need, we work on improving the ergonomics of the system in two ways:

first, ameliorating the reactivity of Coq in its interaction with the user; second, designing and implementing extensions of its interface to ease our formalization activity. We expect the outcome of these lines of research to be useful to a wider audience, interested in manipulating large formulas on topics possibly unrelated to special functions.

2.2.1.3. Formal-proof-producing normalization algorithms

Our algorithm certifications inside Coq intend to simulate well-identified components of our Maple packages, possibly by reproducing them in Coq. It would however not have been judicious to re-implement them inside Coq in a systematic way. Indeed for a number of its components, the output of the algorithm is more easily checked than found, like for instance the solving of a linear system. Rather, we delegate the discovery of the solutions to an external, untrusted oracle like Maple. Trusted computations inside Coq then formally validate the correctness of the a priori untrusted output. More often than not, this validation consists in implementing and executing normalization procedures *inside* Coq. A challenge of this automation is to make sure they go to scale while remaining efficient, which requires a Coq version of non-trivial computer-algebra algorithms. A first, archetypal example we expect to work on is a non-commutative generalization of the normalization procedure for elements of rings [38].

2.2.2. Better symbolic computations with special functions

Generally speaking, we design algorithms for manipulating special functions symbolically, whether univariate or with parameters, and for extracting algorithmically any kind of algebraic and analytic information from them, notably asymptotic properties. Beyond this, the heart of our research is concerned with parametrised definite summations and integrations. These very expressive operations have far-ranging applications, for instance, to the computation of integral transforms (Laplace, Fourier) or to the solution of combinatorial problems expressed via integrals (coefficient extractions, diagonals). The algorithms that we design for them need to really operate on the level of linear functional systems, differential and of recurrence. In all cases, we strive to design our algorithms with the constant goal of good theoretical complexity, and we observe that our algorithms are also fast in practice.

2.2.2.1. Special-function integration and summation

Our long-term goal is to design fast algorithms for a general method for special-function integration (*creative telescoping*), and make them applicable to general special-function inputs. Still, our strategy is to proceed with simpler, more specific classes first (rational functions, then algebraic functions, hyperexponential functions, D-finite functions, non-D-finite functions; two variables, then many variables); as well, we isolate analytic questions by first considering types of integration with a more purely algebraic flavor (constant terms, algebraic residues, diagonals of combinatorics). In particular, we expect to extend our recent approach [22] to more general classes (algebraic with nested radicals, for example): the idea is to speed up calculations by making use of an analogue of Hermite reduction that avoids considering certificates. Homologous problems for summation will be addressed as well.

2.2.2.2. Applications to experimental mathematics

As a consequence of our complexity-driven approach to algorithms design, the algorithms mentioned in the previous paragraph are of good complexity. Therefore, they naturally help us deal with applications that involve equations of high orders and large sizes.

With regard to combinatorics, we expect to advance the algorithmic classification of combinatorial classes like walks and urns. Here, the goal is to determine if enumerative generating functions are rational, algebraic, or D-finite, for example. Physical problems whose modelling involves special-function integrals comprise the study of models of statistical mechanics, like the Ising model for ferro-magnetism, or questions related to Hamiltonian systems.

Number theory is another promising domain of applications. Here, we attempt an experimental approach to the automated certification of integrality of the coefficients of mirror maps for Calabi–Yau manifolds. This could also involve the discovery of new Calabi–Yau operators and the certification of the existing ones. We also plan to algorithmically discover and certify new recurrences yielding good approximants needed in irrationality proofs.

It is to be noted that in all of these application domains, we would so far use general algorithms, as was done in earlier works of ours [21], [25], [24]. To push the scale of applications further, we plan to consider in each case the specifics of the application domain to tailor our algorithms.

2.2.3. Interactive and certified mathematical web sites

In continuation of our past project of an encyclopedia at http://ddmf.msr-inria.inria.fr/1.9.1/ddmf, we ambition to both enrich and certify the formulas about the special functions that we provide online. For each function, our website shows its essential properties and the mathematical objects attached to it, which are often infinite in nature (numerical evaluations, asymptotic expansions). An interactive presentation has the advantage of allowing for adaption to the user's needs. More advanced content will broaden the encyclopedia:

- the algorithmic discussion of equations with parameters, leading to certified automatic case analysis based on arithmetic properties of the parameters;
- lists of summation and integral formulas involving special functions, including validity conditions on the parameters;
- guaranteed large-precision numerical evaluations.

3. Research Program

3.1. Studying special functions by computer algebra

Computer algebra manipulates symbolic representations of exact mathematical objects in a computer, in order to perform computations and operations like simplifying expressions and solving equations for "closed-form expressions". The manipulations are often fundamentally of algebraic nature, even when the ultimate goal is analytic. The issue of efficiency is a particular one in computer algebra, owing to the extreme swell of the intermediate values during calculations.

Our view on the domain is that research on the algorithmic manipulation of special functions is anchored between two paradigms:

- adopting linear differential equations as the right data structure for special functions,
- designing efficient algorithms in a complexity-driven way.

It aims at four kinds of algorithmic goals:

- algorithms combining functions,
- functional equations solving,
- multi-precision numerical evaluations,
- guessing heuristics.

This interacts with three domains of research:

- computer algebra, meant as the search for quasi-optimal algorithms for exact algebraic objects,
- symbolic analysis/algebraic analysis;
- experimental mathematics (combinatorics, mathematical physics, ...).

This view is made explicit in the present section.

3.1.1. Equations as a data structure

Numerous special functions satisfy linear differential and/or recurrence equations. Under a mild technical condition, the existence of such equations induces a finiteness property that makes the main properties of the functions decidable. We thus speak of *D-finite functions*. For example, 60 % of the chapters in the handbook [11] describe D-finite functions. In addition, the class is closed under a rich set of algebraic operations. This makes linear functional equations just the right data structure to encode and manipulate special functions. The power of this representation was observed in the early 1990s [62], leading to the design of many algorithms in computer algebra. Both on the theoretical and algorithmic sides, the study of D-finite functions shares much with neighbouring mathematical domains: differential algebra, D-module theory, differential Galois theory, as well as their counterparts for recurrence equations.

3.1.2. Algorithms combining functions

Differential/recurrence equations that define special functions can be recombined [62] to define: additions and products of special functions; compositions of special functions; integrals and sums involving special functions. Zeilberger's fast algorithm for obtaining recurrences satisfied by parametrised binomial sums was developed in the early 1990s already [63]. It is the basis of all modern definite summation and integration algorithms. The theory was made fully rigorous and algorithmic in later works, mostly by a group in RISC (Linz, Austria) and by members of the team [51], [59], [28], [26], [27], [46]. The past ÉPI Algorithms contributed several implementations (*gfun* [54], *Mgfun* [28]).

3.1.3. Solving functional equations

Encoding special functions as defining linear functional equations postpones some of the difficulty of the problems to a delayed solving of equations. But at the same time, solving (for special classes of functions) is a sub-task of many algorithms on special functions, especially so when solving in terms of polynomial or rational functions. A lot of work has been done in this direction in the 1990s; more intensively since the 2000s, solving differential and recurrence equations in terms of special functions has also been investigated.

3.1.4. Multi-precision numerical evaluation

A major conceptual and algorithmic difference exists for numerical calculations between data structures that fit on a machine word and data structures of arbitrary length, that is, *multi-precision* arithmetic. When multi-precision floating-point numbers became available, early works on the evaluation of special functions were just promising that "most" digits in the output were correct, and performed by heuristically increasing precision during intermediate calculations, without intended rigour. The original theory has evolved in a twofold way since the 1990s: by making computable all constants hidden in asymptotic approximations, it became possible to guarantee a *prescribed* absolute precision; by employing state-of-the-art algorithms on polynomials, matrices, etc, it became possible to have evaluation algorithms in a time complexity that is linear in the output size, with a constant that is not more than a few units. On the implementation side, several original works exist, one of which (*NumGfun* [50]) is used in our DDMF.

3.1.5. Guessing heuristics

"Differential approximation", or "Guessing", is an operation to get an ODE likely to be satisfied by a given approximate series expansion of an unknown function. This has been used at least since the 1970s and is a key stone in spectacular applications in experimental mathematics [25]. All this is based on subtle algorithms for Hermite–Padé approximants [15]. Moreover, guessing can at times be complemented by proven quantitative results that turn the heuristics into an algorithm [23]. This is a promising algorithmic approach that deserves more attention than it has received so far.

3.1.6. Complexity-driven design of algorithms

The main concern of computer algebra has long been to prove the feasibility of a given problem, that is, to show the existence of an algorithmic solution for it. However, with the advent of faster and faster computers, complexity results have ceased to be of theoretical interest only. Nowadays, a large track of works in computer algebra is interested in developing fast algorithms, with time complexity as close as possible to linear in their output size. After most of the more pervasive objects like integers, polynomials, and matrices have been endowed with fast algorithms for the main operations on them [33], the community, including ourselves, started to turn its attention to differential and recurrence objects in the 2000s. The subject is still not as developed as in the commutative case, and a major challenge remains to understand the combinatorics behind summation and integration. On the methodological side, several paradigms occur repeatedly in fast algorithms: "divide and conquer" to balance calculations, "evaluation and interpolation" to avoid intermediate swell of data, etc. [20].

3.2. Trusted computer-algebra calculations

3.2.1. Encyclopedias

Handbooks collecting mathematical properties aim at serving as reference, therefore trusted, documents. The decision of several authors or maintainers of such knowledge bases to move from paper books [11], [13], [55] to websites and wikis [0] allows for a more collaborative effort in proof reading. Another step toward further confidence is to manage to generate the content of an encyclopedia by computer-algebra programs, as is the case with the Wolfram Functions Site [0] or DDMF [0]. Yet, due to the lingering doubts about computer-algebra systems, some encyclopedias propose both cross-checking by different systems and handwritten companion paper proofs of their content. As of today, there is no encyclopedia certified with formal proofs.

3.2.2. Computer algebra and symbolic logic

Several attempts have been made in order to extend existing computer-algebra systems with symbolic manipulations of logical formulas. Yet, these works are more about extending the expressivity of computer-algebra systems than about improving the standards of correctness and semantics of the systems. Conversely, several projects have addressed the communication of a proof system with a computer-algebra system, resulting in an increased automation available in the proof system, to the price of the uncertainty of the computations performed by this oracle.

3.2.3. Certifying systems for computer algebra

More ambitious projects have tried to design a new computer-algebra system providing an environment where the user could both program efficiently and elaborate formal and machine-checked proofs of correctness, by calling a general-purpose proof assistant like the Coq system. This approach requires a huge manpower and a daunting effort in order to re-implement a complete computer-algebra system, as well as the libraries of formal mathematics required by such formal proofs.

3.2.4. Semantics for computer algebra

The move to machine-checked proofs of the mathematical correctness of the output of computer-algebra implementations demands a prior clarification about the often implicit assumptions on which the presumably correctly implemented algorithms rely. Interestingly, this preliminary work, which could be considered as independent from a formal certification project, is seldom precise or even available in the literature.

3.2.5. Formal proofs for symbolic components of computer-algebra systems

A number of authors have investigated ways to organize the communication of a chosen computer-algebra system with a chosen proof assistant in order to certify specific components of the computer-algebra systems, experimenting various combinations of systems and various formats for mathematical exchanges. Another line of research consists in the implementation and certification of computer-algebra algorithms inside the logic [58], [38], [47] or as a proof-automation strategy. Normalization algorithms are of special interest when they allow to check results possibly obtained by an external computer-algebra oracle [31]. A discussion about the systematic separation of the search for a solution and the checking of the solution is already clearly outlined in [44].

3.2.6. Formal proofs for numerical components of computer-algebra systems

Significant progress has been made in the certification of numerical applications by formal proofs. Libraries formalizing and implementing floating-point arithmetic as well as large numbers and arbitrary-precision arithmetic are available. These libraries are used to certify floating-point programs, implementations of mathematical functions and for applications like hybrid systems.

[0] for instance http://dlmf.nist.gov/ for special functions or http://oeis.org/ for integer sequences
[0] http://functions.wolfram.com/
[0] http://ddmf.msr-inria.inria.fr/1.9.1/ddmf

3.3. Machine-checked proofs of formalized mathematics

To be checked by a machine, a proof needs to be expressed in a constrained, relatively simple formal language. Proof assistants provide facilities to write proofs in such languages. But, as merely writing, even in a formal language, does not constitute a formal proof just per se, proof assistants also provide a proof checker: a small and well-understood piece of software in charge of verifying the correctness of arbitrarily large proofs. The gap between the low-level formal language a machine can check and the sophistication of an average page of mathematics is conspicuous and unavoidable. Proof assistants try to bridge this gap by offering facilities, like notations or automation, to support convenient formalization methodologies. Indeed, many aspects, from the logical foundation to the user interface, play an important role in the feasibility of formalized mathematics inside a proof assistant.

3.3.1. Logical foundations and proof assistants

While many logical foundations for mathematics have been proposed, studied, and implemented, type theory is the one that has been more successfully employed to formalize mathematics, to the notable exception of the Mizar system [48], which is based on set theory. In particular, the calculus of construction (CoC) [29] and its extension with inductive types (CIC) [30], have been studied for more than 20 years and been implemented by several independent tools (like Lego, Matita, and Agda). Its reference implementation, Coq [56], has been used for several large-scale formalizations projects (formal certification of a compiler back-end; four-color theorem). Improving the type theory underlying the Coq system remains an active area of research. Other systems based on different type theories do exist and, whilst being more oriented toward software verification, have been also used to verify results of mainstream mathematics (prime-number theorem; Kepler conjecture).

3.3.2. Computations in formal proofs

The most distinguishing feature of CoC is that computation is promoted to the status of rigorous logical argument. Moreover, in its extension CIC, we can recognize the key ingredients of a functional programming language like inductive types, pattern matching, and recursive functions. Indeed, one can program effectively inside tools based on CIC like Coq. This possibility has paved the way to many effective formalization techniques that were essential to the most impressive formalizations made in CIC.

Another milestone in the promotion of the computations-as-proofs feature of Coq has been the integration of compilation techniques in the system to speed up evaluation. Coq can now run realistic programs in the logic, and hence easily incorporates calculations into proofs that demand heavy computational steps.

Because of their different choice for the underlying logic, other proof assistants have to simulate computations outside the formal system, and indeed fewer attempts to formalize mathematical proofs involving heavy calculations have been made in these tools. The only notable exception, which was finished in 2014, the Kepler conjecture, required a significant work to optimize the rewriting engine that simulates evaluation in Isabelle/HOL.

3.3.3. Large-scale computations for proofs inside the Coq system

Programs run and proved correct inside the logic are especially useful for the conception of automated decision procedures. To this end, inductive types are used as an internal language for the description of mathematical objects by their syntax, thus enabling programs to reason and compute by case analysis and recursion on symbolic expressions.

The output of complex and optimized programs external to the proof assistant can also be stamped with a formal proof of correctness when their result is easier to *check* than to *find*. In that case one can benefit from their efficiency without compromising the level of confidence on their output at the price of writing and certify a checker inside the logic. This approach, which has been successfully used in various contexts, is very relevant to the present research project.

3.3.4. Relevant contributions from the Mathematical Component libraries

Representing abstract algebra in a proof assistant has been studied for long. The libraries developed by the MathComp project for the proof of the Odd Order Theorem provide a rather comprehensive hierarchy of structures; however, they originally feature a large number of instances of structures that they need to organize. On the methodological side, this hierarchy is an incarnation of an original work [32] based on various mechanisms, primarily type inference, typically employed in the area of programming languages. A large amount of information that is implicit in handwritten proofs, and that must become explicit at formalization time, can be systematically recovered following this methodology.

Small-scale reflection [35] is another methodology promoted by the MathComp project. Its ultimate goal is to ease formal proofs by systematically dealing with as many bureaucratic steps as possible, by automated computation. For instance, as opposed to the style advocated by Coq's standard library, decidable predicates are systematically represented using computable boolean functions: comparison on integers is expressed as program, and to state that $a \leq b$ one compares the output of this program run on a and b with $true$. In many cases, for example when a and b are values, one can prove or disprove the inequality by pure computation.

The MathComp library was consistently designed after uniform principles of software engineering. These principles range from simple ones, like naming conventions, to more advanced ones, like generic programming, resulting in a robust and reusable collection of formal mathematical components. This large body of formalized mathematics covers a broad panel of algebraic theories, including of course advanced topics of finite group theory, but also linear algebra, commutative algebra, Galois theory, and representation theory. We refer the interested reader to the online documentation of these libraries [57], which represent about 150,000 lines of code and include roughly 4,000 definitions and 13,000 theorems.

Topics not addressed by these libraries and that might be relevant to the present project include real analysis and differential equations. The most advanced work of formalization on these domains is available in the HOL-Light system [40], [41], [42], although some existing developments of interest [18], [49] are also available for Coq. Another aspect of the MathComp libraries that needs improvement, owing to the size of the data we manipulate, is the connection with efficient data structures and implementations, which only starts to be explored.

3.3.5. User interaction with the proof assistant

The user of a proof assistant describes the proof he wants to formalize in the system using a textual language. Depending on the peculiarities of the formal system and the applicative domain, different proof languages have been developed. Some proof assistants promote the use of a declarative language, when the Coq and Matita systems are more oriented toward a procedural style.

The development of the large, consistent body of MathComp libraries has prompted the need to design an alternative and coherent language extension for the Coq proof assistant [37], [36], enforcing the robustness of proof scripts to the numerous changes induced by code refactoring and enhancing the support for the methodology of small-scale reflection.

The development of large libraries is quite a novelty for the Coq system. In particular any long-term development process requires the iteration of many refactoring steps and very little support is provided by most proof assistants, with the notable exception of Mizar [53]. For the Coq system, this is an active area of research.

4. Application Domains

4.1. Computer Algebra in Mathematics

Our expertise in computer algebra and complexity-driven design of algebraic algorithms has applications in various domains, including:

- combinatorics, especially the study of combinatorial walks,
- theoretical computer science, like by the study of automatic sequences,
- number theory, by the analysis of the nature of so-called periods.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

Georges Gonthier, Martìn Abadi and Cédric Fournet receiver the 20 year test-of-time award for their LICS 1998 paper *Secure Implementation of Channel Abstractions*, during LICS 2018 in Oxford.

6. New Software and Platforms

6.1. DynaMoW

Dynamic Mathematics on the Web
FUNCTIONAL DESCRIPTION: Programming tool for controlling the generation of mathematical websites that embed dynamical mathematical contents generated by computer-algebra calculations. Implemented in OCaml.

- Participants: Alexis Darrasse, Frédéric Chyzak and Maxence Guesdon
- Contact: Frédéric Chyzak
- URL: http://ddmf.msr-inria.inria.fr/DynaMoW/

6.2. ECS

Encyclopedia of Combinatorial Structures
FUNCTIONAL DESCRIPTION: On-line mathematical encyclopedia with an emphasis on sequences that arise in the context of decomposable combinatorial structures, with the possibility to search by the first terms in the sequence, keyword, generating function, or closed form.

- Participants: Alexis Darrasse, Frédéric Chyzak, Maxence Guesdon and Stéphanie Petit
- Contact: Frédéric Chyzak
- URL: http://ecs.inria.fr/

6.3. DDMF

Dynamic Dictionary of Mathematical Functions
FUNCTIONAL DESCRIPTION: Web site consisting of interactive tables of mathematical formulas on elementary and special functions. The formulas are automatically generated by OCaml and computer-algebra routines. Users can ask for more terms of the expansions, more digits of the numerical values, proofs of some of the formulas, etc.

- Participants: Alexandre Benoit, Alexis Darrasse, Bruno Salvy, Christoph Koutschan, Frédéric Chyzak, Marc Mezzarobba, Maxence Guesdon, Stefan Gerhold and Thomas Gregoire
- Contact: Frédéric Chyzak
- URL: http://ddmf.msr-inria.inria.fr/1.9.1/ddmf

6.4. Mgfun

multivariate generating functions package

FUNCTIONAL DESCRIPTION: The Mgfun Project is a collection of packages for the computer algebra system Maple, and is intended for the symbolic manipulation of a large class of special functions and combinatorial sequences (in one or several variables and indices) that appear in many branches of mathematics, mathematical physics, and engineering sciences. Members of the class satisfy a crucial finiteness property which makes the class amenable to computer algebra methods and enjoy numerous algorithmic closure properties, including algorithmic closures under integration and summation.

- Contact: Frédéric Chyzak
- URL: http://specfun.inria.fr/chyzak/mgfun.html

6.5. Ssreflect

FUNCTIONAL DESCRIPTION: Ssreflect is a tactic language extension to the Coq system, developed by the Mathematical Components team.

- Participants: Assia Mahboubi, Cyril Cohen, Enrico Tassi, Georges Gonthier, Laurence Rideau, Laurent Théry and Yves Bertot
- Contact: Yves Bertot
- URL: http://math-comp.github.io/math-comp/

6.6. Math-Components

Mathematical Components library
FUNCTIONAL DESCRIPTION: The Mathematical Components library is a set of Coq libraries that cover the mechanization of the proof of the Odd Order Theorem.

RELEASE FUNCTIONAL DESCRIPTION: The library includes 16 more theory files, covering in particular field and Galois theory, advanced character theory, and a construction of algebraic numbers.

- Participants: Alexey Solovyev, Andrea Asperti, Assia Mahboubi, Cyril Cohen, Enrico Tassi, François Garillot, Georges Gonthier, Ioana Pasca, Jeremy Avigad, Laurence Rideau, Laurent Théry, Russell O'Connor, Sidi Ould Biha, Stéphane Le Roux and Yves Bertot
- Contact: Assia Mahboubi
- URL: http://math-comp.github.io/math-comp/

7. New Results

7.1. Computing solutions of linear Mahler equations

Mahler equations relate evaluations of the same function f at iterated bth powers of the variable. They arise in particular in the study of automatic sequences and in the complexity analysis of divide-and-conquer algorithms. Recently, the problem of solving Mahler equations in closed form has occurred in connection with number-theoretic questions. A difficulty in the manipulation of Mahler equations is the exponential blow-up of degrees when applying a Mahler operator to a polynomial. In [3], Frédéric Chyzak and Philippe Dumas, together with Thomas Dreyfus (IRMA, Université de Strasbourg) and Marc Mezzarobba (external collaborator from Sorbonne Université), have presented algorithms for solving linear Mahler equations for series, polynomials, and rational functions, and have obtained polynomial-time complexity under a mild assumption. The article was formally accepted and published this year.

7.2. Becker's conjecture on Mahler functions

In 1994, Becker conjectured that if $F(z)$ is a k-regular power series, then there exists a k-regular rational function $R(z)$ such that $F(z)/R(z)$ satisfies a Mahler-type functional equation with polynomial coefficients where the initial coefficient satisfies $a_0(z) = 1$. In [1], Frédéric Chyzak and Philippe Dumas, together with Jason P. Bell (University of Waterloo, Canada) and Michael Coons (University of Newcastle, Australia) have proved Becker's conjecture in the best-possible form: they have shown that the rational function $R(z)$ can be taken to be a polynomial $z^\gamma Q(z)$ for some explicit non-negative integer γ and such that $1/Q(z)$ is k-regular. The article was formally accepted this year.

7.3. Generalized Hermite reduction, creative telescoping and definite integration of D-finite functions

Hermite reduction is a classical algorithmic tool in symbolic integration. It is used to decompose a given rational function as a sum of a function with simple poles and the derivative of another rational function. Alin Bostan, Frédéric Chyzak, and Pierre Lairez, together with Bruno Salvy (project-team AriC) have extended Hermite reduction to arbitrary linear differential operators instead of the pure derivative. They have also developped efficient algorithms for this reduction, and then applied the generalized Hermite reduction to the computation of linear operators satisfied by single definite integrals of D-finite functions of several continuous or discrete parameters. The resulting algorithm is a generalization of reduction-based methods for creative telescoping. Their article [6] was published at the ISSAC conference.

7.4. Bijections between Łukasiewicz walks and generalized tandem walks

In [9], Frédéric Chyzak, together with Karen Yeats (University of Waterloo, Canada), have studied the enumeration by length of several walk models on the square lattice. They have obtained bijections between walks in the upper half-plane returning to the x-axis and walks in the quarter plane. An ongoing work by Bostan, Chyzak, and Mahboubi has given a bijection for models using small north, west, and south-east steps. The work in [9] has adapted and generalized it to a bijection between half-plane walks using those three steps in two colours and a quarter-plane model over the symmetrized step set consisting of north, north-west, west, south, south-east, and east. They have then generalized their bijections to certain models with large steps: for given $p \geq 1$, a bijection has been given between the half-plane and quarter-plane models obtained by keeping the small south-east step and replacing the two steps north and west of length 1 by the $p + 1$ steps of length p in directions between north and west. An article was submitted this year.

7.5. Putting Fürer's algorithm into practice with the BPAS library

Fast algorithms for integer and polynomial multiplication play an important role in scientific computing as well as in other disciplines. In 1971, Schönhage and Strassen designed an algorithm that improved the multiplication time for two integers of at most n bits to $O(\log n \log \log n)$. Martin Fürer presented a new algorithm that runs in $O(n \log n \cdot 2^{O(\log^* n)})$, where $\log^* n$ is the iterated logarithm of n. In a submitted article, Svyatoslav Covanov, together with Davood Mohajerani, Marc Moreno Maza and Lin-Xiao Wang, have explained how one can put Fürer's ideas into practice for multiplying polynomials over a prime field $\mathbb{Z}/p\mathbb{Z}$, for which p is a Generalized Fermat prime of the form $p = r^k + 1$ where k is a power of 2 and r is of machine word size. When k is at least 8, they have shown that multiplication inside such a prime field can be efficiently implemented via Fast Fourier Transform (FFT). Taking advantage of Cooley-Tukey tensor formula and the fact that r is a $2k$-th primitive root of unity in $\mathbb{Z}/p\mathbb{Z}$, they have obtained an efficient implementation of FFT over $\mathbb{Z}/p\mathbb{Z}$. This implementation outperforms comparable implementations either using other encodings of $\mathbb{Z}/p\mathbb{Z}$ or other ways to perform multiplication in $\mathbb{Z}/p\mathbb{Z}$.

7.6. Fast coefficient computation for algebraic power series in positive characteristic

In [5], Alin Bostan and Philippe Dumas, together with Xavier Caruso (CNRS, Rennes) and Gilles Christol (IMJ, Paris) have studied the algorithmic question of coefficient computation of algebraic power series in positive characteristic. They revisited Christol's theorem on algebraic power series in positive characteristic and proposed another proof for it. Their new proof combines several ingredients and advantages of existing proofs, which make it very well-suited for algorithmic purposes. The construction used in the new proof was then applied to the design of a new efficient algorithm for computing the Nth coefficient of a given algebraic power series over a perfect field of characteristic p. This algorithm has several nice features: it is more general, more natural and more efficient than previous algorithms. Not only the arithmetic complexity of the new algorithm is linear in $\log N$ and quasi-linear in p, but its dependency with respect to the degree of the input is much smaller than in the previously best algorithm. Moreover, when the ground field is finite, the new approach yields an even faster algorithm, whose bit complexity is linear in $\log N$ and quasi-linear in $\sqrt{p}$.

7.7. Counting walks with large steps in an orthant

In the past fifteen years, the enumeration of lattice walks with steps taken in a prescribed set and confined to a given cone, especially the first quadrant of the plane, has been intensely studied. As a result, the generating functions of quadrant walks are now well-understood, provided the allowed steps are *small*. In particular, having small steps is crucial for the definition of a certain group of bi-rational transformations of the plane. It has been proved that this group is finite if and only if the corresponding generating function is D-finite. This group is also the key to the uniform solution of 19 of the 23 small step models possessing a finite group. In contrast, almost nothing was known for walks with arbitrary steps. In [7], Alin Bostan together with Mireille Bousquet-Mélou (CNRS, Bordeaux) and Stephen Melczer (U. Pennsylvania, Philadelphia, USA), extended the definition of the group, or rather of the associated orbit, to this general case, and generalized the above uniform solution of small step models. When this approach works, it invariably yields a D-finite generating function. They applied it to many quadrant problems, including some infinite families. After developing the general theory, the authors of [7] considered the 13 110 two-dimensional models with steps in $\{-2,-1,0,1\}^2$ having at least one -2 coordinate. They proved that only 240 of them have a finite orbit, and solve 231 of them with our method. The 9 remaining models are the counterparts of the 4 models of the small step case that resist the uniform solution method (and which are known to have an algebraic generating function). They conjecture D-finiteness for their generating functions (but only two of them are likely to be algebraic!), and proved non-D-finiteness for the 12 870 models with an infinite orbit, except for 16 of them.

7.8. Subresultants of $(x-\alpha)^m$ and $(x-\beta)^n$, Jacobi polynomials and complexity

A previous article in 2017 described explicit expressions for the coefficients of the order-d polynomial subresultant of $(x-\alpha)^m$ and $(x-\beta)^n$ with respect to Bernstein's set of polynomials $\{(x-\alpha)^j(x-\beta)^{d-j},\ 0\le j\le d\}$, for $0\le d<\min\{m,n\}$. In [8], Alin Bostan, together with T. Krick, M. Valdettaro (U. Buenos Aires, Argentina) and A. Szanto (U. North Carolina, Raleigh, USA) further developed the study of these structured polynomials and showed that the coefficients of the subresultants of $(x-\alpha)^m$ and $(x-\beta)^n$ with respect to the monomial basis can be computed in *linear* arithmetic complexity, which is faster than for arbitrary polynomials. The result is obtained as a consequence of the amazing though seemingly unnoticed fact that these subresultants are scalar multiples of Jacobi polynomials up to an affine change of variables.

7.9. A numerical transcendental method in algebraic geometry

In "A transcendental method in algebraic geometry", Griffiths emphasized the role of certain multivariate integrals, known as *periods*, "to construct a continuous invariant of arbitrary smooth projective varieties".

Periods often determine the projective variety completely and therefore its algebraic invariants. Translating periods into discrete algebraic invariants is a difficult problem, exemplified by the long standing Hodge conjecture which describes how periods determine the algebraic cycles within a projective variety.

Recent progress in computer algebra makes it possible to compute periods with high precision and put transcendental methods into practice. In [10], Pierre Lairez and Emre Sertöz focus on algebraic surfaces and give a numerical method to compute Picard groups. As an application, they count smooth rational curves on quartic surfaces using the Picard group. It is the first time that this kind of computation is performed.

8. Partnerships and Cooperations

8.1. International Research Visitors

8.1.1. Internships

- Jiadong Han did a Master internship from March to August. Under the supervision of Pierre Lairez, he studied the computation of adaptive grid to improve the computation of the homology of semialgebraic sets.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

- Alin Bostan is part of the Scientific advisory board of the conference series *Effective Methods in Algebraic Geometry* (MEGA).
- Alin Bostan was a member of the Scientific advisory board of the conference *Algèbre, arithmétique et combinatoire des équations différentielles et aux différences*, CIRM (Luminy, France); ~ 60 participants.
- Alin Bostan is part of the scientific committee of the GDR EFI (“Functional Equations and Interactions”) dependent on the mathematical institute (INSMI) of the CNRS. The goal of this GDR is to bring together various research communities in France working on functional equations in fields of computer science and mathematics.
- Frédéric Chyzak is member of the steering committee of the *Journées Nationales de Calcul Formel* (JNCF), the annual meeting of the French computer algebra community.
- Frédéric Chyzak is elected member (and current chair) of the steering committee of the *International Symposium on Symbolic and Algebraic Computation* (ISSAC, 3-year term, 2016–2018).
- Georges Gonthier is a member of the steering committee of the *Certified Programs and Proofs* Conference (CPP).

9.1.1.1. Member of the Organizing Committees

- Alin Bostan co-organizes, with Lucia Di Vizio, the *Séminaire Différentiel* between U. Versailles and Inria Saclay, with a bi-annual frequency (~ 30 participants per event).
- Alin Bostan co-organizes, with Lucia Di Vizio, the working group *Marches dans le quart de plan*, at Institut Henri Poincaré (Paris), with a bi-monthly frequency (~ 15 participants per event).

9.1.2. Scientific Events Selection

9.1.2.1. Reviewer

- Frédéric Chyzak has served as reviewer for the selection of the international conference ISSAC 2018.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- Alin Bostan is on the editorial board of the *Journal of Symbolic Computation*.
- Georges Gonthier is on the editorial board of the *Journal of Formalized Reasoning*.

9.1.3.2. Reviewer - Reviewing Activities

- Alin Bostan has served as a reviewer for the journals: *Journal of Symbolic Computation*, *Journal of Combinatorial Theory, Series A*, *Applicable Algebra in Engineering Communications and Computing*, *Minnesota Journal of Undergraduate Mathematics*.
- Frédéric Chyzak has served multiple times as a reviewer for the *Journal of Symbolic Computation*.
- Pierre Lairez has served as a reviewer for the *Journal of Symbolic Computation*, *Journal of the ACM* and *Journal of Physics A*.

9.1.4. Invited Talks

- Alin Bostan has been invited to give a talk at the *Workshop on algebraic and analytic aspects of power series*, Universidade Lisboa, Lisbonne, Portugal, Jan. 2018.
- Alin Bostan has been invited to give a talk at the conference *Algebra, Arithmetic and Combinatorics of Differential and Difference Equations*, CIRM (Luminy), France, May 2018.
- Alin Bostan has been invited to give a talk at the conference *Grands réseaux aléatoires et marches contraintes*, in honor of the 75th birthday of Guy Fayolle, Dijon, France, Aug. 2018.
- Alin Bostan has been invited to give a talk at the conference *Combinatorics and Arithmetic for Physics: special days*, IHES, Bures-sur-Yvette, France, Oct. 2018.
- Frédéric Chyzak was invited invited speaker at the conference *Rencontres Arithmétiques du GDR Informatique Mathématique* (RAIM 2018), Gif-sur-Yvette, France, Nov. 2018.
- Georges Gonthier was a plenary keynote speaker at the *Federated Logic Conference (FLoC 2018)* in Oxford, July 2018.
- Georges Gonthier was invited speaker at the *Workshop on Modular Knowledge (Tetrapod)* during FLoC 2018, Oxford, July 2018.
- Georges Gonthier ws the keynote speaker of the *Future of Mathematical Proofs* workshop at the Heidelberg Laureate Forum, September 2018.

9.1.5. Leadership within the Scientific Community

9.1.5.1. Regular Research Seminar

The team organizes a regular seminar, with roughly 15–20 talks a year. The topics reflect the team's interests: computer algebra, combinatorics, number theory, formal proofs, and related domains. This year, we reduced a bit the number of talks in our seminar, as we have invested much time in setting up a working group with a talk every second week (see 9.1.5.2).

9.1.5.2. Research Working Group

This year we have set up a working group *Marches dans le quart de plan* around the study of walks in the quarter plan, a very active research topic in probability theory and enumerative combinatorics in recent years. The working group is organized at Institut Henri Poincaré, with a regularity of two sessions per month. The original purpose was to read the article "On the Nature of the Generating Series of Walks in the Quarter Plane" by T. Dreyfus, C. Hardouin, J. Roques, M. Singer, published in Invent. Math. this year. But the reality exceeded expectations: the working group attracted a dozen of people, working either in computer science or pure mathematics, who began to interact and a very good dynamic was created. Altogether, sixteen sessions have taken place so far, and we have decided to continue in 2019. From the team, Alin Bostan, Frédéric Chyzak, Guy Fayolle, and Pierre Lairez have given a total of 9 talks to this working group.

9.1.6. Scientific Expertise

- Georges Gonthier participated in a review of the software and algorithms of the Tezos blockchain conducted by the Inria Foundation during Spring 2018.

9.1.7. Research Administration

- Georges Gonthier serves on the Conseil de l'École Doctorale de Mathématiques Hadamard.

9.2. Teaching - Supervision - Juries

- Alin Bostan has served as a jury member of the French *Agrégation de Mathématiques – épreuve de modélisation, option C*.

9.2.1. Teaching

Licence:

Pierre Lairez, *Introduction à l'informatique (INF311)*, TD, 40h, L3, École polytechnique, France.

Master:

Frédéric Chyzak, *Algorithmes efficaces en calcul formel*, 18h, M2, MPRI, France.

Alin Bostan, *Algorithmes efficaces en calcul formel*, 40.5h, M2, MPRI, France.

Pierre Lairez, *Algorithmique avancée (INF550)*, TD, 18h, M2, École polytechnique, France.

Pierre Lairez, *Les bases de la programmation et de l'algorithmique (INF411)*, TD, 40h, M1, École polytechnique, France.

9.2.2. Juries

- Frédéric Chyzak has been a member of the hiring jury at Inria (Concours CRCN 2018).
- Alin Bostan has served as a referee in the PhD jury of Timothée Pecatte, *Bornes inférieures et algorithmes de reconstruction pour des sommes de puissances affines*, ENS Lyon, July 11, 2018.
- Alin Bostan has served as an examiner in the PhD jury of Boris Djalal, *Formalisations en Coq pour la décision de problèmes en géométrie algébrique réelle*, Inria Sophia Antipolis, December 3, 2018.
- Alin Bostan has served as a member of the monitoring PhD committee of Youssef Abdelaziz, Univ. Paris 6.
- Alin Bostan has served as a member of the monitoring PhD committee of Manon Bertin, Univ. Rouen.

9.3. Popularization

9.3.1. Interventions

- Georges Gonthier testified before the *Mission d'information commune sur les blockchains* of the *Assemblée Nationale* in March.
- Georges Gonthier gave a public lecture and debate on blockchains at the *Institut Diderot* in September, jointly with M. Odonnat (Banque de France).

9.3.2. Internal action

- Georges Gonthier gave a presentation at the *Journées Scientifiques Inria 2018* in Bordeaux.

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] J. P. BELL, F. CHYZAK, M. COONS, P. DUMAS.*Becker's conjecture on Mahler functions*, in "Transactions of the American Mathematical Society", 2018, 17, In press, https://hal.inria.fr/hal-01885598

[2] P. BÜRGISSER, F. CUCKER, P. LAIREZ.*Computing the Homology of Basic Semialgebraic Sets in Weak Exponential Time*, in "Journal of the ACM (JACM)", December 2018, vol. 66, n^{o} 1, p. 1-30 [*DOI :* 10.1145/3275242], https://hal.archives-ouvertes.fr/hal-01545657

[3] F. CHYZAK, T. DREYFUS, P. DUMAS, M. MEZZAROBBA.*Computing solutions of linear Mahler equations*, in "Mathematics of Computation", July 2018, vol. 87, p. 2977-3021 [*DOI :* 10.1090/MCOM/3359], https://hal.inria.fr/hal-01418653

[4] A. MAHBOUBI, G. MELQUIOND, T. SIBUT-PINOTE.*Formally Verified Approximations of Definite Integrals*, in "Journal of Automated Reasoning", March 2018, p. 1-20 [*DOI :* 10.1007/S10817-018-9463-7], https://hal.inria.fr/hal-01630143

International Conferences with Proceedings

[5] A. BOSTAN, X. CARUSO, G. CHRISTOL, P. DUMAS.*Fast Coefficient Computation for Algebraic Power Series in Positive Characteristic*, in "Thirteenth Algorithmic Number Theory Symposium ANTS-XIII", Madison, United States, Algorithmic Number Theory, July 2018, https://arxiv.org/abs/1806.06543 , https://hal.archives-ouvertes.fr/hal-01816375

[6] A. BOSTAN, F. CHYZAK, P. LAIREZ, B. SALVY.*Generalized Hermite Reduction, Creative Telescoping and Definite Integration of D-Finite Functions*, in "ISSAC 2018 - International Symposium on Symbolic and Algebraic Computation", New York, United States, July 2018, p. 1-8 [*DOI :* 10.1145/3208976.3208992], https://hal.inria.fr/hal-01788619

Other Publications

[7] A. BOSTAN, M. BOUSQUET-MÉLOU, S. MELCZER.*Counting walks with large steps in an orthant*, May 2018, https://arxiv.org/abs/1806.00968 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01802706

[8] A. BOSTAN, T. KRICK, A. SZANTO, M. VALDETTARO.*Subresultants of $(x-\alpha)^m$ and $(x-\beta)^n$, Jacobi polynomials and complexity*, December 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01966640

[9] F. CHYZAK, K. YEATS.*Bijections between Łukasiewicz walks and generalized tandem walks*, October 2018, working paper or preprint, https://hal.inria.fr/hal-01891792

[10] P. LAIREZ, E. CAN SERTÖZ.*A numerical transcendental method in algebraic geometry*, November 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01932147

References in notes

[11] M. ABRAMOWITZ, I. A. STEGUN (editors). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, Dover, New York, 1992, xiv+1046, Reprint of the 1972 edition

[12] *Computer Algebra Errors*, Article in mathematics blog MathOverflow, http://mathoverflow.net/questions/11517/computer-algebra-errors

[13] F. W. J. OLVER, D. W. LOZIER, R. F. BOISVERT, C. W. CLARK (editors). *NIST Handbook of mathematical functions*, Cambridge University Press, 2010

[14] M. ARMAND, B. GRÉGOIRE, A. SPIWACK, L. THÉRY.*Extending Coq with Imperative Features and its Application to SAT Verication*, in "Interactive Theorem Proving, international Conference, ITP 2010, Edinburgh, Scotland, July 11–14, 2010, Proceedings", Lecture Notes in Computer Science, Springer, 2010

[15] B. BECKERMANN, G. LABAHN.*A uniform approach for the fast computation of matrix-type Padé approximants*, in "SIAM J. Matrix Anal. Appl.", 1994, vol. 15, n^{o} 3, p. 804–823

[16] A. BENOIT, F. CHYZAK, A. DARRASSE, S. GERHOLD, M. MEZZAROBBA, B. SALVY.*The Dynamic Dictionary of Mathematical Functions (DDMF)*, in "The Third International Congress on Mathematical Software (ICMS 2010)", K. FUKUDA, J. VAN DER HOEVEN, M. JOSWIG, N. TAKAYAMA (editors), Lecture Notes in Computer Science, 2010, vol. 6327, p. 35–41, http://dx.doi.org/10.1007/978-3-642-15582-6_7

[17] M. BOESPFLUG, M. DÉNÈS, B. GRÉGOIRE.*Full reduction at full throttle*, in "First International Conference on Certified Programs and Proofs, Taiwan, December 7–9", Lecture Notes in Computer Science, Springer, 2011

[18] S. BOLDO, C. LELAY, G. MELQUIOND.*Improving Real Analysis in Coq: A User-Friendly Approach to Integrals and Derivatives*, in "Certified Programs and Proofs", C. HAWBLITZEL, D. MILLER (editors), Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2012, vol. 7679, p. 289-304, http://dx.doi.org/10.1007/978-3-642-35308-6_22

[19] S. BOLDO, G. MELQUIOND.*Flocq: A Unified Library for Proving Floating-point Algorithms in Coq*, in "Proceedings of the 20th IEEE Symposium on Computer Arithmetic", Tübingen, Germany, July 2011, p. 243–252

[20] A. BOSTAN.*Algorithmes rapides pour les polynômes, séries formelles et matrices*, in "Actes des Journées Nationales de Calcul Formel", Luminy, France, 2010, p. 75–262, Les cours du CIRM, tome 1, numéro 2, http://ccirm.cedram.org:80/ccirm-bin/fitem?id=CCIRM_2010__1_2_75_0

[21] A. BOSTAN, S. BOUKRAA, S. HASSANI, J.-M. MAILLARD, J.-A. WEIL, N. ZENINE.*Globally nilpotent differential operators and the square Ising model*, in "J. Phys. A: Math. Theor.", 2009, vol. 42, n^{o} 12, 50, http://dx.doi.org/10.1088/1751-8113/42/12/125206

[22] A. BOSTAN, S. CHEN, F. CHYZAK, Z. LI.*Complexity of creative telescoping for bivariate rational functions*, in "ISSAC'10: Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation", New York, NY, USA, ACM, 2010, p. 203–210, http://doi.acm.org/10.1145/1837934.1837975

[23] A. BOSTAN, F. CHYZAK, G. LECERF, B. SALVY, É. SCHOST.*Differential equations for algebraic functions*, in "ISSAC'07: Proceedings of the 2007 international symposium on Symbolic and algebraic computation", C. W. BROWN (editor), ACM Press, 2007, p. 25–32, http://dx.doi.org/10.1145/1277548.1277553

[24] A. BOSTAN, F. CHYZAK, M. VAN HOEIJ, L. PECH.*Explicit formula for the generating series of diagonal 3D rook paths*, in "Sém. Loth. Comb.", 2011, vol. B66a, 27, http://www.emis.de/journals/SLC/wpapers/s66bochhope.html

[25] A. BOSTAN, M. KAUERS.*The complete generating function for Gessel walks is algebraic*, in "Proceedings of the American Mathematical Society", September 2010, vol. 138, n^{o} 9, p. 3063–3078, With an appendix by Mark van Hoeij

[26] F. CHYZAK.*An extension of Zeilberger's fast algorithm to general holonomic functions*, in "Discrete Math.", 2000, vol. 217, n° 1-3, p. 115–134, Formal power series and algebraic combinatorics (Vienna, 1997)

[27] F. CHYZAK, M. KAUERS, B. SALVY.*A Non-Holonomic Systems Approach to Special Function Identities*, in "ISSAC'09: Proceedings of the Twenty-Second International Symposium on Symbolic and Algebraic Computation", J. MAY (editor), 2009, p. 111–118, http://dx.doi.org/10.1145/1576702.1576720

[28] F. CHYZAK, B. SALVY.*Non-commutative elimination in Ore algebras proves multivariate identities*, in "J. Symbolic Comput.", 1998, vol. 26, n° 2, p. 187–227

[29] T. COQUAND, G. P. HUET.*The Calculus of Constructions*, in "Inf. Comput.", 1988, vol. 76, n° 2/3, p. 95-120, http://dx.doi.org/10.1016/0890-5401(88)90005-3

[30] T. COQUAND, C. PAULIN-MOHRING.*Inductively defined types*, in "Proceedings of Colog'88", P. MARTIN-LÖF, G. MINTS (editors), Lecture Notes in Computer Science, Springer-Verlag, 1990, vol. 417

[31] D. DELAHAYE, M. MAYERO.*Dealing with algebraic expressions over a field in Coq using Maple*, in "J. Symbolic Comput.", 2005, vol. 39, n° 5, p. 569–592, Special issue on the integration of automated reasoning and computer algebra systems, http://dx.doi.org/10.1016/j.jsc.2004.12.004

[32] F. GARILLOT, G. GONTHIER, A. MAHBOUBI, L. RIDEAU.*Packaging Mathematical Structures*, in "Theorem Proving in Higher-Order Logics", S. BERGHOFER, T. NIPKOW, C. URBAN, M. WENZEL (editors), Lecture Notes in Computer Science, Springer, 2009, vol. 5674, p. 327–342

[33] J. VON ZUR. GATHEN, J. GERHARD.*Modern computer algebra*, 2nd, Cambridge University Press, New York, 2003, xiv+785

[34] G. GONTHIER.*Formal proofs—the four-colour theorem*, in "Notices of the AMS", 2008, vol. 55, n° 11, p. 1382-1393

[35] G. GONTHIER, A. MAHBOUBI.*An introduction to small scale reflection in Coq*, in "Journal of Formalized Reasoning", 2010, vol. 3, n° 2, p. 95–152

[36] G. GONTHIER, A. MAHBOUBI, E. TASSI.*A Small Scale Reflection Extension for the Coq system*, Inria, 2008, n° RR-6455, http://hal.inria.fr/inria-00258384

[37] G. GONTHIER, E. TASSI.*A language of patterns for subterm selection*, in "ITP", LNCS, 2012, vol. 7406, p. 361–376

[38] B. GRÉGOIRE, A. MAHBOUBI.*Proving Equalities in a Commutative Ring Done Right in Coq*, in "Theorem Proving in Higher Order Logics, 18th International Conference, TPHOLs 2005, Oxford, UK, August 22-25, 2005, Proceedings", Lecture Notes in Computer Science, Springer, 2005, vol. 3603, p. 98–113

[39] T. HALES.*Formal proof*, in "Notices of the AMS", 2008, vol. 55, n° 11, p. 1370-1380

[40] J. HARRISON.*A HOL Theory of Euclidean space*, in "Theorem Proving in Higher Order Logics, 18th International Conference, TPHOLs 2005", Oxford, UK, J. HURD, T. MELHAM (editors), Lecture Notes in Computer Science, Springer-Verlag, 2005, vol. 3603

[41] J. HARRISON.*Formalizing an analytic proof of the prime number theorem*, in "Journal of Automated Reasoning", 2009, vol. 43, p. 243–261, Dedicated to Mike Gordon on the occasion of his 60th birthday

[42] J. HARRISON.*Theorem proving with the real numbers*, CPHC/BCS distinguished dissertations, Springer, 1998

[43] J. HARRISON.*A Machine-Checked Theory of Floating Point Arithmetic*, in "Theorem Proving in Higher Order Logics: 12th International Conference, TPHOLs'99", Nice, France, Y. BERTOT, G. DOWEK, A. HIRSCHOWITZ, C. PAULIN, L. THÉRY (editors), Lecture Notes in Computer Science, Springer-Verlag, 1999, vol. 1690, p. 113–130

[44] J. HARRISON, L. THÉRY.*A Skeptic's Approach to Combining HOL and Maple*, in "J. Autom. Reason.", December 1998, vol. 21, n^{o} 3, p. 279–294, http://dx.doi.org/10.1023/A:1006023127567

[45] F. JOHANSSON.*Another Mathematica bug*, 2009, Article on personal blog, http://fredrik-j.blogspot.fr/2009/07/another-mathematica-bug.html

[46] C. KOUTSCHAN.*A fast approach to creative telescoping*, in "Math. Comput. Sci.", 2010, vol. 4, n^{o} 2-3, p. 259–266, http://dx.doi.org/10.1007/s11786-010-0055-0

[47] A. MAHBOUBI.*Implementing the cylindrical algebraic decomposition within the Coq system*, in "Mathematical Structures in Computer Science", 2007, vol. 17, n^{o} 1, p. 99–127

[48] R. MATUSZEWSKI, P. RUDNICKI.*Mizar: the first 30 years*, in "Mechanized Mathematics and Its Applications", 2005, vol. 4

[49] M. MAYERO.*Problèmes critiques et preuves formelles*, Université Paris 13, novembre 2012, Habilitation à Diriger des Recherches

[50] M. MEZZAROBBA.*NumGfun: a package for numerical and analytic computation and D-finite functions*, in "ISSAC 2010—Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation", New York, ACM, 2010, p. 139–146, http://dx.doi.org/10.1145/1837934.1837965

[51] P. PAULE, M. SCHORN.*A Mathematica version of Zeilberger's algorithm for proving binomial coefficient identities*, in "J. Symbolic Comput.", 1995, vol. 20, n^{o} 5-6, p. 673–698, Symbolic computation in combinatorics Δ_1 (Ithaca, NY, 1993), http://dx.doi.org/10.1006/jsco.1995.1071

[52] B. PETERSEN.*Maple*, Personal web site

[53] P. RUDNICKI, A. TRYBULEC.*On the Integrity of a Repository of Formalized Mathematics*, in "Proceedings of the Second International Conference on Mathematical Knowledge Management", London, UK, MKM '03, Springer-Verlag, 2003, p. 162–174, http://dl.acm.org/citation.cfm?id=648071.748518

[54] B. SALVY, P. ZIMMERMANN.*Gfun: a Maple package for the manipulation of generating and holonomic functions in one variable*, in "ACM Trans. Math. Software", 1994, vol. 20, n^{o} 2, p. 163–177

[55] N. J. A. SLOANE, S. PLOUFFE.*The Encyclopedia of Integer Sequences*, Academic Press, San Diego, 1995

[56] THE COQ DEVELOPMENT TEAM.*The Coq Proof Assistant: Reference Manual*, http://coq.inria.fr/doc/

[57] THE MATHEMATICAL COMPONENT TEAM.*A Formalization of the Odd Order Theorem using the Coq proof assistant*, September 2012, http://www.msr-inria.fr/projects/mathematical-components/

[58] L. THÉRY.*A Machine-Checked Implementation of Buchberger's Algorithm*, in "J. Autom. Reasoning", 2001, vol. 26, n^{o} 2, p. 107-137, http://dx.doi.org/10.1023/A:1026518331905

[59] K. WEGSCHAIDER.*Computer generated proofs of binomial multi-sum identities*, RISC, J. Kepler University, May 1997, 99

[60] S. WOLFRAM.*Mathematica: A system for doing mathematics by computer (2nd ed.)*, Addison-Wesley, 1992

[61] D. ZEILBERGER.*Opinion 94: The Human Obsession With "Formal Proofs" is a Waste of the Computer's Time, and, Even More Regretfully, of Humans' Time*, 2009, http://www.math.rutgers.edu/~zeilberg/Opinion94.html

[62] D. ZEILBERGER.*A holonomic systems approach to special functions identities*, in "J. Comput. Appl. Math.", 1990, vol. 32, n^{o} 3, p. 321–368

[63] D. ZEILBERGER.*The method of creative telescoping*, in "J. Symbolic Comput.", 1991, vol. 11, n^{o} 3, p. 195–204

Team TAU

TAckling the Underspecified

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

RESEARCH CENTER
Saclay - Île-de-France

THEME
Optimization, machine learning and statistical methods

Table of contents

Team TAU

Creation of the Team: 2016 December 01

Keywords:

Computer Science and Digital Science:

A3.3.3. - Big data analysis
A3.4. - Machine learning and statistics
A3.5.2. - Recommendation systems
A8.2. - Optimization
A8.6. - Information theory
A9.2. - Machine learning
A9.3. - Signal analysis

Other Research Topics and Application Domains:

B1.1.4. - Genetics and genomics
B4. - Energy
B7.2.1. - Smart vehicles
B9.1.2. - Serious games
B9.5.3. - Physics
B9.5.6. - Data science
B9.6.10. - Digital humanities

1. Team, Visitors, External Collaborators

Research Scientists

Marc Schoenauer [Team leader, Inria, Senior Researcher, HDR]
Guillaume Charpiat [Inria, Researcher]
Cyril Furtlehner [Inria, Researcher]
Flora Jay [CNRS, Researcher]
Michèle Sebag [CNRS, Senior Researcher, HDR]
Paola Tubaro [CNRS, Researcher]

Faculty Members

Cécile Germain [Univ Paris-Sud, Emeritus]
Nicolas Spyratos [Univ Paris-Sud, Emeritus]
Philippe Caillou [Univ Paris-Sud, Associate Professor]
Aurélien Decelle [Univ Paris-Sud, Associate Professor]
Isabelle Guyon [Univ Paris-Sud, Professor; Inria/UPSaclay Data Science Chair]
Francois Landes [Univ Paris-Sud, Associate Professor, from Sep 2018]
Véronique Ventos [Univ Paris-Sud, Associate Professor, until August 2018]

External Collaborator

Yann Ollivier [Facebook AI Research, HDR]

Technical Staff

Laurent Basara [Inria, from Nov 2018]
Jonas Renault [Inria, from Nov 2018]
Yasmina Bouzbiba [Inria, until Mar 2018]
Etienne Brame [Inria]

Raphael Jaiswal [Inria]
Felix Louistisserand [Inria, until Jan 2018]

PhD Students

Eleonore Bartenlian [Univ Paris-Sud, from Oct 2018]
Victor Berger [Inria]
Leonard Blier [Facebook, from Sep 2018]
Benjamin Donnot [RTE]
Balthazar Donon [RTE, from Oct. 2018]
Guillaume Doquet [Univ Paris-Sud]
Victor Estrade [Univ Paris-Sud]
Loris Felardos Saint Jean [Inria, from Oct 2018]
Giancarlo Fissore [Univ Paris-Sud]
Julien Girard [CEA, from Oct 2018]
François Gonard [Institut de recherche technologique System X, until Jun 2018]
Diviyan Kalainathan [Univ Paris-Sud]
Zhengying Liu [Ecole polytechnique]
Nizam Makdoud [Thales, from Feb 2018]
Marc Nabhan [Renault]
Adrian Pol [Organisation européenne pour la recherche nucléaire]
Herilalaina Rakotoarison [Inria]
Théophile Sanchez [Univ Paris-Sud]
Thomas Schmitt [Univ Paris-Sud]
Lisheng Sun [Univ Paris-Sud]
Corentin Tallec [Ecole polytechnique]
Marion Ullmo [CNRS, from Oct 2018]
Pierre Wolinski [Ecole Normale Supérieure Paris]
Mandar Chandorkar [CWI]

Post-Doctoral Fellows

Olivier Goudet [Inria, until Aug. 2018]
Nayat Sanchez Pi [Inria, from Jun 2018 until Aug 2018]
Berna Bakir Batu [Inria]

Administrative Assistant

Laurence Fontana [Inria]

2. Overall Objectives

2.1. Presentation

The last two years have been a turning point for the team. Of course, since its creation in 2003, TAO activities had constantly but slowly evolved, as old problems were being solved, and new applications arose. But recent abrupt progresses in Machine Learning (and in particular in Deep Learning) have greatly accelerated these changes also within the team. It so happened that this change of slope also coincided with some more practical changes in TAO ecosystem: following Inria 12-years rule, the team definitely ended in December 2016. The new team TAU (for **TA**ckling the **U**nderspecified) has been proposed, and the creation process is on-going. At the same time important staff changes took place, that also justify even sharper changes in the team focus. During the year 2018, the second year of this new era for the (remaining) members of the team, our research topics have now stabilized around a final version of the TAU project.

Following the dramatic changes in TAU staff during the years 2016-2017 (see the 2017 activity report of the team for the details), the research around continuous optimization has definitely faded out in TAU (while the research axis on hyperparameter tuning has focused on Machine Learning algorithms), the Energy application domain has slightly changed direction under Isabelle Guyon's supervision (Section 4.2), after the completion of the work started by Olivier Teytaud, and a few new directions have emerged, though yet too young to have given any visible fruit, around the robustness of ML systems (Section 3.1.2). The other research topics have been continued, as described below.

2.2. Context and overall goal of the project

Building upon the expertise in machine learning (ML) and optimization of the TAO team, the TAU project will tackle some **under-specified challenges behind the New Artificial Intelligence wave**. The simultaneous advent of massive data and massive computational power, blurring the boundaries between data, structure, knowledge and common sense, seemingly makes it possible to fulfill all promises of the good old AI, now or soon.

This makes NewAI under-specified in three respects. A first dimension regards the relationships between AIs and human beings. The necessary conditions for AIs to be accepted by mankind and/or contribute to the common good are yet to be formally defined; it is hard to believe that a general and computable definition of "ethical behavior" can be set once for all. Some of these necessary conditions (explainable and causal modeling; unbiased data and models; model certification) can nevertheless be cast as ambitious and realistic goals for public research.

A second dimension regards the relationships between AI, data and knowledge. In closed worlds AIs can manage and acquire sufficient data to reach human-level performances from scratch [81]. In open worlds however, prior knowledge is used in various ways to overcome the lack of direct interactions with the world, e.g. through i) exploiting domain-dependent data invariances in intension or in extension (ranging from convolution to domain augmentation); ii) taking advantage of the low-rank structure (generative learning) or known properties (equivariant learning) of the observed data; iii) leveraging diverse domains and datasets, assumedly related to each other (domain adaptation; multi-task learning). A general and open question is how available prior knowledge can be best leveraged by an AI, all the more so as domains with small to medium-size data are considered.

A third dimension regards the intrinsic limitations of AI in terms of information theory. Long established theories, e.g. rooted in Occam's razor, currently hardly account for the practical leaps of deep learning, where the solution dimension outnumbers the input dimension. Beyond trials-and-errors, a long-term goal is to characterize the learning landscape w.r.t. order parameters to be defined, and *a priori* estimate the regions of problem instances where it is likely/possible/unlikely to learn accurate models.

The above under-specified AI issues define three core research pillars (Section 3), examining three interdependent aspects of AI:

I. The first pillar aims to answer the question of what it means to be a good AI and how to build them. More specifically, our goal is to advance the state of the art concerning robust learning (re adversarial attacks), causal modeling (aimed to support explanations and prescriptions), and unbiased models in the sense of prescribed neutrality constraints (including the assessment and repair of the data).

II. The second pillar tackles the "innate *vs* acquired" question: how to best combine available human knowledge, and agnostic machine learning. TAU will examine this question focusing on domains with spatial and temporal multi-scale structure, as pervasive in natural sciences (where domain knowledge is expressed using PDEs, or through powerful compact representations as in signal processing), taking advantage of the pluri-disciplinary expertise and scientific collaborations of the TAU members.

III. The third pillar aims to understand the learning landscape. In the short term, it tackles the so-called Auto-☆ issue of automatically selecting and configuring an algorithm portfolio for a problem instance. This issue governs the knowledge transfer from research labs to industry [92], [91], all the more so as massive computational resources are at stake. In the medium term, our goal is to integrate the hyper-parameters and model structure in the learning criteria, using information theory and/or bilevel programming [93]. In the long-term, our goal is to establish a phase diagram of the learning landscape, through i) determining order parameters; ii) relating the different regions defined along these order parameters, to the quality of the optimal solution, and the probability of finding a good approximation thereof. These goals are aligned with the unique scientific expertise of TAU in statistical physics and in information theory, and benefit from our decade-long expertise in Auto-☆.

The above research pillars will take inspiration and be validated with three applicative topics (Section 4):

1.Energy management encompasses a variety of scientific problems related to research pillars I. (fair learning, privacy-compliant modelling, safety-related guarantees) and II. (spatio-temporal multi-scale modelling, distributional learning). It is also a strategic application for the planet, where TAU benefits from the TAO expertise and the long established relationships with Artelys (ILab Metis) and RTE.

2.Computational Social Sciences offer questions and methodological lessons about how to address these questions in a common decency spirit, along research pillar I. On-going studies at TAU include the learning and randomized assessment of prescriptive models for Human Resources (hiring and vocational studies; quality of life at work and economic performance) and nutrition habits (in relation with social networks and health), where i) learned models must be unbiased although data are undoubtedly biased; ii) prior knowledge must be accounted for and the interpretation of the learned models is mandatory; iii) causal modelling is key as models are deployed for *prescription* and self-fulfilling prophecies must be avoided at all costs [131].

3.Optimal data-driven design considers several physical/simulated phenomenons, ranging from high-energy physics to space weather, from population biology to medical imaging, from signal processing to certification of autonomous vehicle controllers, with: i) medium-size data; ii) extensive prior knowledge, notably concerning the symmetries and properties of the sought models; iii) computationally expensive simulators. All three characteristics are relevant to pillars II and III.

3. Research Program

3.1. Toward Good AI

As discussed by [134], the topic of ethical AI was non-existent until 2010, was laughed at in 2016, and became a hot topic in 2017 as the AI disruptivity with respect to the fabric of life (travel, education, entertainment, social networks, politics, to name a few) became unescapable [131], together with its expected impacts on the nature and amount of jobs. As of now, it seems that the risk of a new AI Winter might arise from legal[0] and societal[0] issues. While privacy is now recognized as a civil right in Europe, it is feared that the GAFAM, BATX and others can already capture a sufficient fraction of human preferences and their dynamics to achieve their commercial and other goals, and build a Brave New Big Brother (BNBB, a system that is openly beneficial to many, covertly nudging, and possibly dictatorial).

The ambition of TAU is to mitigate the BNBB risk along several intricated dimensions, and build i) causal and explainable models; ii) fair data and models; iii) provably robust models.

3.1.1. Causal modeling and biases

Participants: Isabelle Guyon, Michèle Sebag, Philippe Caillou, Paola Tubaro
PhD: Diviyan Kalainathan
Collaboration: Olivier Goudet (TAU then Univ. Angers), David Lopez-Paz (Facebook)

The extraction of causal models, a long goal of AI [132], [112], [133], became a strategic issue as the usage of learned models gradually shifted from *prediction* to *prescription* in the last years. This evolution, following Auguste Comte's vision of science (*Savoir pour prévoir, afin de pouvoir*) indeed reflects the exuberant optimism about AI: Knowledge enables Prediction; Prediction enables Control. However, although predictive models can be based on correlations, prescriptions can only be based on causal models[0].

[0]For instance, the (fictitious) plea challenge proposed to law students in Oct. 2018 considered a chain reaction pileup occurred among autonomous and humanly operated vehicles on a highway.

[0]For instance related to information bubbles and nudge [98], [148].

Among the research applications concerned with causal modeling, predictive modeling or collaborative filtering at TAU are all projects described in section 4.1 (see also Section 3.4), studying the relationships between: i) the educational background of persons and the job openings (FUI project JobAgile and DataIA project Vadore); ii) the quality of life at work and the economic performance indicators of the enterprises (ISN Lidex project Amiqap) [114] ; iii) the nutritional items bought by households (at the level of granularity of the barcode) and their health status, as approximated from their body-mass-index (IRS UPSaclay Nutriperso); iv) the actual offer of restaurants and their scores on online rating systems. In these projects, a wealth of data is available (though hardly sufficient for applications ii), iii and iv))) and there is little doubt that these data reflect the imbalances and biases of the world as is, ranging from gender to racial to economical prejudices. Preventing the learned models from perpetuating such biases is essential to deliver an AI endowed with common decency.

In some cases, the bias is known; for instance, the cohorts in the Nutriperso study are more well-off than the average French population, and the Kantar database includes explicit weights to address this bias through importance sampling. In other cases, the bias is only guessed; for instance, the companies for which Secafi data are available hardly correspond to a uniform sample as these data have been gathered upon the request of the company trade union.

3.1.2. Robustness of Learned Models

Participants: Guillaume Charpiat, Marc Schoenauer, Michèle Sebag
PhD and Engineers: Julien Girard, Marc Nabhan, Nizham Makhoud, Raphaël Jaiswal
Collaboration: Zakarian Chihani (CEA); Hiba Hage, Philippe Reynaud, and Yves Tourbier (Renault)

Due to their outstanding performances, deep neural networks and more generally machine learning-based decision making systems, referred to as MLs in the following, have been raising hopes in the recent years to achieve breakthroughs in critical systems, ranging from autonomous vehicles to defense. The main pitfall for such applications lies in the lack of guarantees for MLs robustness.

Specifically, MLs are used when the mainstream software design process does not apply, that is, when no formal specification of the target software behavior is available and/or when the system is embedded in an open unpredictable world. The extensive body of knowledge developed to deliver guarantees about mainstream software – ranging from formal verification, model checking and abstract interpretation to testing, simulation and monitoring – thus does not directly apply either. Another weakness of MLs regards their dependency to the amount and quality of the training data, as their performances are sensitive to slight perturbations of the data distribution. Such perturbations can occur naturally due to domain or concept drift (e.g. due to a change in light intensity or a scratch on a camera lens); they can also result from intentional malicious attacks, a.k.a adversarial examples [149].

These downsides, currently preventing the dissemination of MLs in safety-critical systems (SCS), call for a considerable amount of research, in order to understand when and to which extent an MLs can be certified to provide the desired level of guarantees.

Julien Girard's PhD (CEA scholarship), started in Oct. 2018, co-supervised by Guillaume Charpiat and Zakaria Chihani (CEA), is devoted to the extension of abstract interpretation to deep neural nets, and the formal characterization of the transition kernel from input to output space achieved by a DNN (robustness by design, coupled with formally assessing the coverage of the training set). This approach is tightly related to the inspection and opening of black-box models, aimed to characterize the patterns in the input instances responsible for a decision – another step toward explainability.

[0]One can predict that it rains based on the presence of umbrellas in the street; but one cannot induce rainfall by going out with an umbrella. Likewise, the presence of books/tablets at home and the good scores of children at school are correlated; but offering books/tablets to all children might fail to improve their scores *per se*, if both good scores and books are explained by a so-called confounder variable, like the presence of adults versed in books/tablets at home.

On the other hand, experimental validation of MLs, akin statistical testing, also faces three limitations: i) real-world examples are notoriously insufficient to ensure a good coverage in general; ii) for this reason, simulated examples are extensively used; but their use raises the *reality gap* issue [123] of the distance between real and simulated worlds; iii) independently, the real-world is naturally subject to domain shift (e.g. due to the technical improvement and/or aging of sensors). Our collaborations with Renault tackle such issues in the context of the autonomous vehicle (see Section 7.1.3).

3.2. Hybridizing numerical modeling and learning systems

Participants: Guillaume Charpiat, Cécile Germain, Isabelle Guyon, Marc Schoenauer, Michèle Sebag
PhD: Théophile Sanchez, Loris Felardo
In sciences and engineering, human knowledge is commonly expressed in closed form, through equations or mechanistic models characterizing how a natural or social phenomenon, or a physical device, will behave/evolve depending on its environment and external stimuli, under some assumptions and up to some approximations. The field of numerical engineering, and the simulators based on such mechanistic models, are at the core of most approaches to understand and analyze the world, from solid mechanics to computational fluid dynamics, from chemistry to molecular biology, from astronomy to population dynamics, from epidemiology and information propagation in social networks to economy and finance.

Most generally, numerical engineering supports the simulation, and when appropriate the optimization and control[0] of the phenomenons under study, although several sources of discrepancy might adversely affect the results, ranging from the underlying assumptions and simplifying hypotheses in the models, to systematic experiment errors to statistical measurement errors (not to mention numerical issues). This knowledge and know-how are materialized in millions of lines of code, capitalizing the expertise of academic and industrial labs. These softwares have been steadily extended over decades, modeling new and more fine-grained effects through layered extensions, making them increasingly harder to maintain, extend and master. Another difficulty is that complex systems most often resort to hybrid (pluridisciplinary) models, as they involve many components interacting along several time and space scales, hampering their numerical simulation.

At the other extreme, machine learning offers the opportunity to model phenomenons from scratch, using any available data gathered through experiments or simulations. Recent successes of machine learning in computer vision, natural language processing and games to name a few, have demonstrated the power of such agnostic approaches and their efficiency in terms of prediction [118], inverse problem solving [161], and sequential decision making [154], [81], despite their lack of any "semantic" understanding of the universe. Even before these successes, Anderson's claim was that *the data deluge [might make] the scientific method obsolete* [69], as if a reasonable option might be to throw away the existing equational or software bodies of knowledge, and let Machine Learning rediscover all models from scratch. Such a claim is hampered among others by the fact that not all domains offer a wealth of data, as any academic involved in an industrial collaboration around data has discovered.

Another approach will be considered in TAU, investigating how existing mechanistic models and related simulators can be partnered with ML algorithms: i) to achieve the same goals with the same methods with a gain of accuracy or time; ii) to achieve new goals; iii) to achieve the same goals with new methods.

Toward more robust numerical engineering: In domains where satisfying mechanistic models and simulators are available, ML can contribute to improve their accuracy or usability. A first direction is to refine or extend the models and simulators to better fit the empirical evidence. The goal is to finely account for the different biases and uncertainties attached to the available knowledge and data, distinguishing the different types of *known unknowns*. Such *known unknowns* include the model hyper-parameters (coefficients), the systematic errors due to e.g., experiment imperfections, and the statistical errors due to e.g., measurement errors. A second approach is based on learning a surrogate model for the phenomenon under study that incorporate domain knowledge from the mechanistic model (or its simulation). See Section 7.5 for case studies.

[0]Note that the causal nature of mechanistic models is established from prior knowledge and experimentations.

A related direction, typically when considering black-box simulators, aims to learn a model of the error, or equivalently, a post-processor of the software. The discrepancy between simulated and empirical results, referred to as *reality gap* [123], can be tackled in terms of domain adaptation [74], [97]. Specifically, the source domain here corresponds to the simulated phenomenon, offering a wealth of inexpensive data, and the target domain corresponds to the actual phenomenon, with rare and expensive data; the goal is to devise accurate target models using the source data and models.

Extending numerical engineering: ML, using both experimental and numerical data, can also be used to tackle new goals, that are beyond the current state-of-the-art of standard approaches. Inverse problems are such goals, identifying the parameters or the initial conditions of phenomenons for which the model is not differentiable, or amenable to the adjoint state method.

A slightly different kind of inverse problem is that of recovering the ground truth when only noisy data is available. This problem can be formulated as a search for the simplest model explaining the data. The question then becomes to formulate and efficiently exploit such a simplicity criterion.

Another goal can be to model the distribution of given quantiles for some system: The challenge is to exploit available data to train a generative model, aimed at sampling the target quantiles.

Examples tackled in TAU are detailed in Section 7.5. Note that the "Cracking the Glass Problem", described in Section 7.2.3 is yet another instance of a similar problem.

Data-driven numerical engineering : Finally, ML can also be used to sidestep numerical engineering limitations in terms of scalability, or to build a simulator emulating the resolution of the (unknown) mechanistic model from data, or to revisit the formal background.

When the mechanistic model is known and sufficiently accurate, it can be used to train a deep network on an arbitrary set of (space,time) samples, resulting in a meshless numerical approximation of the model [145], supporting by construction *differentiable programming* [120].

When no mechanistic model is sufficiently efficient, the model must be identified from the data only. Genetic programming has been used to identify systems of ODEs [142], through the identification of invariant quantities from data, as well as for the direct identification of control commands of nonlinear complex systems, including some chaotic systems [85]. Another recent approach uses two deep neural networks, one for the state of the system, the other for the equation itself [135]. The critical issues for both approaches include the scalability, and the explainability of the resulting models. Such line of research will benefit from TAU unique mixed expertise in Genetic Programming and Deep Learning.

Finally, in the realm of signal processing (SP), the question is whether and how deep networks can be used to revisit mainstream feature extraction based on Fourier decomposition, wavelet and scattering transforms [77]. E. Bartenlian's PhD (started Oct. 2018), co-supervised by M. Sebag and F. Pascal (Centrale-Supélec), focusing on musical audio-to-score translation [144], inspects the effects of supervised training, taking advantage from the fact that convolution masks can be initialized and analyzed in terms of frequency.

3.3. Learning to learn

According to Ali Rahimi's test of times award speech at NIPS 17, the current ML algorithms *have become a form of alchemy*. Competitive testing and empirical breakthroughs gradually become mandatory for a contribution to be acknowledged; an increasing part of the community adopts trials and errors as main scientific methodology, and theory is lagging behind practice. This style of progress is typical of technological and engineering revolutions for some; others ask for consolidated and well-understood theoretical advances, saving the time wasted in trying to build upon hardly reproducible results.

Basically, while practical achievements have often passed the expectations, there exist caveats along three dimensions. Firstly, excellent performances do not imply that the model has captured what was to learn, as shown by the phenomenon of adversarial examples. Following Ian Goodfellow, some well-performing models might be compared to *Clever Hans*, the horse that was able to solve mathematical exercizes using non verbal cues from its teacher [111]; it is the purpose of Pillar I. to alleviate the *Clever Hans* trap (section 3.1).
Secondly, some major advances, e.g. related to the celebrated adversarial learning [101], [97], establish proofs of concept more than a sound methodology, where the reproducibility is limited due to i) the computational power required for training (often beyond reach of academic labs); ii) the numerical instabilities (witnessed as random seeds happen to be found in the codes); iii) the insufficiently documented experimental settings. What works, why and when is still a matter of speculation, although better understanding the limitations of the current state of the art is acknowledged to be a priority. After Ali Rahimi again, *simple experiments, simple theorems are the building blocks that help us understand more complicated systems*. Along this line, [128] propose toy examples to demonstrate and understand the defaults of convergence of gradient descent adversarial learning.
Thirdly, and most importantly, the reported achievements rely on carefully tuned learning architectures and hyper-parameters. The sensitivity of the results to the selection and calibration of algorithms has been identified since the end 80s as a key ML bottleneck, and the field of automatic algorithm selection and calibration, referred to as AutoML or Auto-☆ in the following, is at the ML forefront.

TAU aims to contribute to the ML evolution toward a more mature stage along three dimensions. In the short term, the research done in Auto-☆ will be pursued (section 3.3.1). In the medium term, an information theoretic perspective will be adopted to capture the data structure and to calibrate the learning algorithm *depending on the nature and amount of the available data*. In the longer term, our goal is to leverage the methodologies forged in statistical physics to understand and control the trajectories of complex learning systems (section 3.3.3).

3.3.1. Auto-*

Participants: Isabelle Guyon, Marc Schoenauer, Michèle Sebag
PhD: Guillaume Doquet, Zhengying Liu, Herilalaina Rakotoarison, Lisheng Sun
Collaboration: Olivier Bousquet, André Elisseeff (Google Zurich)

The so-called Auto-☆ task, concerned with selecting a (quasi) optimal algorithm and its hyper-parameters depending on the problem instance at hand, remained a key issue in ML for the last three decades [75], as well as in optimization at large [110], including combinatorial optimization and constraint satisfaction [117], [100] and continuous optimization [71]. This issue, tackled by several European projects along the decades, governs the knowledge transfer to industry, due to the shortage of data scientists. It becomes even more crucial as models are more complex and their training requires more computational resources. This has motivated several international challenges devoted to Auto-ML [45] (see also Section 3.4), including the on-going AutoDL [37] (see also Section 7.6).

Several approaches have been used to tackle Auto-☆ in the literature, and TAU has been particularly active in the first two. Meta-learning aims to build a surrogate performance model, estimating the performance of an algorithm configuration on *any* problem instance characterized from its meta-feature values [138], [100], [72], [71] [11]. Collaborative filtering, considering that a problem instance "likes better" an algorithm configuration yielding a better performance, learns to recommend good algorithms to problem instances [147], [130]. Bayesian optimization proceeds by alternatively building a surrogate model of algorithm performances on *the* problem instance at hand, and tackling it [92]. This last approach currently is the prominent one; as shown in [130], the meta-features developed for AutoML are hardly relevant, hampering both meta-learning and collaborative filtering.

Beyond these, current research directions in TAU include the design of more efficient features, as well as the design of an original approach based on MCTS algorithm (see Section 7.2.1).

3.3.2. Information theory: adjusting model complexity and data fitting

Participants: Guillaume Charpiat, Marc Schoenauer, Michèle Sebag
PhD: Corentin Tallec, Pierre Wolinski, Léonard Blier
Collaboration: Yann Ollivier (Facebook)

In the 60s, Kolmogorov and Solomonoff provided a well-grounded theory for building (probabilistic) models best explaining the available data [139], [103], that is, the shortest programs able to generate these data. Such programs can then be used to generate further data or to answer specific questions (interpreted as missing values in the data). Deep learning, from this viewpoint, efficiently explores a space of computation graphs, described from its hyperparameters (network structure) and parameters (weights). Network training amounts to optimizing these parameters, namely, navigating the space of computational graphs to find a network, as simple as possible, that explain the past observations well.

This vision is at the core of variational auto-encoders [116], directly optimizing a bound on the Kolmogorov complexity of the dataset. More generally variational methods provide quantitative criteria to identify superfluous elements (edges, units) in a neural network, that can potentially be used for structural optimization of the network (Leonard Blier's PhD, started Oct. 2018).

The same principles apply to unsupervised learning, aimed to find the maximum amount of structure hidden in the data, quantified using this information-theoretic criterion.

The known invariances in the data can be exploited to guide the model design (e.g. as translation invariance leads to convolutional structures, or LSTM is shown to enforce the invariance to time affine transformations of the data sequence [150]). Scattering transforms exploit similar principles [77]. A general theory of how to detect *unknown* invariances in the data, however, is currently lacking.

The view of information theory and Kolmogorov complexity suggests that key program operations (composition, recursivity, use of predefined routines) should intervene when searching for a good computation graph. One possible framework for exploring the space of computation graphs with such operations is that of genetic programming [70]. It is interesting to see that evolutionary computation appeared in the last two years among the best candidates to explore the space of deep learning structures [137], [122]. Other approaches might proceed by combining simple models into more powerful ones, e.g. using "Context Tree Weighting" [158] or switch distributions [88]. Another option is to formulate neural architecture design as a reinforcement learning problem [73]; the value of the building blocks (predefined routines) might be defined using e.g., Monte-Carlo Tree Search. A key difficulty is the computational cost of retraining neural nets from scratch upon modifying their architecture; an option might be to use neutral initializations to support warm-restart.

3.3.3. Analyzing and Learning Complex Systems

Participants: Cyril Furtlehner, Aurélien Decelle, François Landes, Michèle Sebag
PhD: Giancarlo Fissore
Collaboration: Enrico Camporeale (CWI); Jacopo Rocchi (LPTMS Paris Sud), the Simons team: Rahul Chako (post-doc), Andrea Liu (UPenn), David Reichman (Columbia), Giulio Biroli (ENS), Olivier Dauchot (ESPCI).

Methods and criteria from statistical physics have been widely used in ML. In early days, the capacity of Hopfield networks (associative memories defined by the attractors of an energy function) was investigated by using the replica formalism [68]. Restricted Boltzmann machines likewise define a generative model built upon an energy function trained from the data. Along the same lines, Variational Auto-Encoders can be interpreted as systems relating the free energy of the distribution, the information about the data and the entropy (the degree of ignorance about the micro-states of the system) [157]. A key promise of the statistical physics perspective and the Bayesian view of deep learning is to harness the tremendous growth of the model size (billions of weights in recent machine translation netwowrks), and make them sustainable through e.g. weight quantization [125], posterior drop-out [129], and probabilistic binary networks [126]. Such "informational cooling" of a trained deep network can reduce its size by several orders of magnitude while preserving its performance.

Statistical physics is among the key expertises of TAU, originally only represented by Cyril Furtlehner, later strenghtened by Aurélien Decelle's and François Landes' arrivals in 2014 and 2018. On-going studies are conducted along several directions.

Generative models are most often expressed in terms of a Gibbs distributions $P[S] = \exp(-E[S])$, where energy E involves a sum of building blocks, modelling the interactions among variables. This formalization makes it natural to use mean-field methods of statistical physics and associated inference algorithms to both train and exploit such models. The difficulty is to find a good trade-off between the richness of the structure and the efficiency of mean-field approaches. One direction of research pursued in TAU, [94] in the context of traffic forecasting, is to account for the presence of cycles in the interaction graph, to adapt inference algorithms to such graphs with cycles, while constraining graphs to remain compatible with mean-field inference.

Another direction, explored in TAO in the recent years, is based on the definition and exploitation of self-consistency properties, enforcing principled divide-and-conquer resolutions. In the particular case of the message-passing Affinity Propagation algorithm for instance [159], self-consistency imposes the invariance of the solution when handled at different scales, thus enabling to characterize the critical value of the penalty and other hyper-parameters in closed form (in the case of simple data distributions) or empirically otherwise [95].

A more recent research direction examines the quantity of information in a (deep) neural net along the random matrix theory framework [80]. It is addressed in Giancarlo Fissore's PhD, and is detailed in Section 7.2.3.

A collaboration with L. Zbederova's group at Ecole Normale Supérieure is just starting, thanks to François Landes' arrival, based on the study of some information-theoretic indicators for some class of Neural Network. It is, too, described in more details in Section 7.2.3.

Finally, we note the recent surge in using ML to address fundamental physics problems: from turbulence to high-energy physics and soft matter as well (with glasses at its core). TAU's dual expertise in Deep Networks and in statistical physics places it in an ideal position to significantly contribute to this domain and shape the methods that will be used by the physics community in the future. François Landes' recent arrival in the team makes TAU a unique place for such interdisciplinary research, thanks to his collaborators from the Simons Collaboration *Cracking the Glass Problem* (gathering 13 statistical physics teams at the international level). This project is detailed in Section 7.2.3.

3.4. Organisation of Challenges

Participants: Cécile Germain, Isabelle Guyon, Marc Schoenauer, Michèle Sebag
Challenges have been an important drive for Machine Learning research for many years, and TAO members have played important roles in the organization of many such challenges: Michèle Sebag was head of the challenge programme in the Pascal European Network of Excellence (2005-2013); Isabelle Guyon, as mentioned, was the PI of many challenges ranging from causation challenges [104], to AutoML [105]. The Higgs challenge [65], most attended ever Kaggle challenge, was jointly organized by TAO (C. Germain), LAL-IN2P3 (D. Rousseau and B. Kegl) and I. Guyon (not yet at TAO), in collaboration with CERN and Imperial College.

TAU was also particularly implicated with the ChaLearn Looking At People (LAP) challenge series in Computer Vision, in collaboration with the University of Barcelona [90] including the Job Candidate Screening Coopetition [89]; the Real Versus Fake Expressed Emotion Challenge (ICCV 2017) [155]; the Large-scale Continuous Gesture Recognition Challenge (ICCV 2017) [155]; the Large-scale Isolated Gesture Recognition Challenge (ICCV 2017) [155].

Other challenges have been organized in 2018, or are planned for the near future, detailed in Section 7.6. In particular, many of them now run on the Codalab platform, managed by TAU and maintained at LRI.

4. Application Domains

4.1. Computational Social Sciences

Participants: Philippe Caillou, Isabelle Guyon, Michèle Sebag, Paola Tubaro
Collaboration: Jean-Pierre Nadal (EHESS); Marco Cuturi, Bruno Crépon (ENSAE); Thierry Weil (Mines); Jean-Luc Bazet (RITM)

Computational Social Sciences (CSS) studies social and economic phenomena, ranging from technological innovation to politics, from media to social networks, from human resources to education, from inequalities to health. It combines perspectives from different scientific disciplines, building upon the tradition of computer simulation and modeling of complex social systems [99] on the one hand, and data science on the other hand, fueled by the capacity to collect and analyze massive amounts of digital data.

The emerging field of CSS raises formidable challenges along three dimensions. Firstly, the definition of the research questions, the formulation of hypotheses and the validation of the results require a tight pluridisciplinary interaction and dialogue between researchers from different backgrounds. Secondly, the development of CSS is a touchstone for ethical AI. On the one hand, CSS gains ground in major, data-rich private companies; on the other hand, public researchers around the world are engaging in an effort to use it for the benefit of society as a whole [119]. The key technical difficulties related to data and model biases, and to self-fulfilling prophecies have been discussed in section 3.1. Thirdly, CSS does not only regard scientists: it is essential that the civil society participate in the science of society [146].

TAO was involved in CSS for the last five years, and its activities have been strengthened thanks to P. Tubaro's and I. Guyon's expertises respectively in sociology and economics, and in causal modeling. Details are given in Section 7.3.

4.2. Energy Management

Participants: Isabelle Guyon, Marc Schoenauer, Michèle Sebag
PhD: Victor Berger, Benjamin Donnot, Balthazar Donon, Herilalaina Rakotoarison
Collaboration: Antoine Marot, Patrick Panciatici (RTE), Vincent Renault (Artelys), Olivier Teytaud (Facebook)

Energy Management has been an application domain of choice for TAO since the end 2000s, with main partners SME Artelys (METIS Ilab Inria; ADEME project POST; ADEME project NEXT) and RTE (See.4C European challenge; two CIFRE PhDs). The goals concern i) optimal planning over several spatio-temporal scales, from investments on continental Europe/North Africa grid at the decade scale (POST), to daily planning of local or regional power networks (NEXT); ii) monitoring and control of the French grid enforcing the prevention of power breaks (RTE); iii) improvement of house-made numerical methods using data-intense learning (as described in Section 3.2) in all aspects of IFPEN activities, from geological problems in oil prospection (IFPEN) to the optimal placement of eolians in eolian fields (IFPEN).
Optimal planning over long periods of time amounts to optimal sequential decision under high uncertainties, ranging from stochastic uncertainties (weather, market prices, demand prediction) handled based on massive data, to non-stochastic uncertainties (e.g., political decisions about the nuclear policy) handled through defining and selecting a tractable number of scenarios. Note that non-anticipativity constraints forbid the use of dynamic programming-related methods; this led to propose the *Direct Value Search* method [79] at the end of the POST project. A further recent work in the same direction [21] proposes and theoretically studies the *Direct Model Predictive Control* approach, a hybrid model which merges the properties of two different dynamic optimization methods, Model Predictive Control and Stochastic Dual Dynamic Programming, has robust convergence properties, and experimentally competes with both methods alone.

The daily maintainance of power grids requires the building of approximate predictive models on the top of any given network topology. Deep Networks are natural candidates for such modelling, considering the size of the French grid ($\sim$ 10000 nodes), but the representation of the topology is a challenge when, e.g. the RTE goal is to quickly ensure the "n-1" security constraint (the network should remain safe even if any of the 10000 nodes fails). Existing simulators are too slow to be used in real time, and the size of actual grids makes it intractable to train surrogate models for all possible (n-1) topologies (see Section 7.4 for more details).

Even when efficient simulators do exist, they need to be calibrated (adjusting their hyper-parameters with real data), and complemented by uncertainty propagation models. Such adaptations and extensions are at the core of the NEXT project; hyper-parameter tuning is also a challenge regarding the development plans of the local grids, that heavily rely on graph optimization algorithms.
Furthermore, predictive models of local grids are based on the estimated consumption of end-customers: Linky meters provide coarse grain information only due to privacy issues, and very few samples of fine-grained consumption are available (from volunteer customers). A first task is to transfer knowledge from small data to the whole domain of application. A second task is to directly predict the peak of consumption based on the user cluster profiles and their representativity (see Section 7.4.2).

Another research direction formulates security maintenance as a reinforcement problem, taking inspiration from the recent successes of Deep Reinforcement Learning. This direction is being investigated in Balthazar Donon's RTE CIFRE PhD with RTE (started Oct. 2018).

4.3. Data-driven Numerical Modeling

Participants: Guillaume Charpiat, Cécile Germain, Isabelle Guyon, Flora Jay, Marc Schoenauer, Michèle Sebag
PhD and Post-doc: Victor Estrade, Loris Felardo, Adrian Pol, Théophile Sanchez
Collaboration: D. Rousseau (LAL), M. Pierini (CERN)

As said (section 3.2), in domains where both first principle-based models and equations, and empirical or simulated data are available, their combined usage can support more accurate modelling and prediction, and when appropriate, optimization, control and design. This section describes such applications, with the goal of improving the time-to-design chain through fast interactions between the simulation, optimization, control and design stages. The expected advances regard: i) the quality of the models or simulators (through data assimilation, e.g. coupling first principles and data, or repairing/extending closed-form models); ii) the exploitation of data derived from different distributions and/or related phenomenons; and, most interestingly, iii) the task of optimal design and the assessment of the resulting designs.

The proposed approaches are based on generative and adversarial modelling [116], [102], extending both the generator and the discriminator modules to take advantage of the domain knowledge.

A first challenge regards the design of the model space, and the architecture used to enforce the known domain properties (symmetries, invariance operators, temporal structures). When appropriate, data from different distributions (e.g. simulated vs real-world data) will be reconciled, for instance taking inspiration from real-valued non-volume preserving transformations [84] in order to preserve the natural interpretation.
Another challenge regards the validation of the models and solutions of the optimal design problems. The more flexible the models, the more intensive the validation must be, as reminded by Leon Bottou. Along this way, generative models will be used to support the design of "what if" scenarios, to enhance anomaly detection and monitoring via refined likelihood criteria.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

- *GECCO 2018 10-years impact award*, awarded to the paper published in GECCO 2008 that had the greatest impact, seen from 10 years later, for the paper

  Adaptive operator selection with dynamic multi-armed bandits, by Luis DaCosta, Alvaro Fialho, Marc Schoenauer, and Michèle Sebag, in Maarten Keijzer (Ed), Proc. ACM-GECCO, pp 913-920, 2008.
- Nacim Belkhir, Winner ACM-GECCO 2018 **BBComp single-objective** and **expensive single-objective** tracks. Nacim completed his PhD in TAU in 2017 [71], co-supervised by Marc Schoenauer, Johann Dréo and Pierre Savéant (Thalès TRT).

5.1.2. Visibility

- Marc Schoenauer, member of the core team responsible for the *Villani mission* regarding the French strategy on Artifical Intelligence. The mission started Sept. 2017 and the final report was delivered on March 29. 2018.
- Michèle Sebag, elected member of French Académie des Technologies, Apr. 2018.
- Michèle Sebag, chevalière de la Légion d'Honneur, Dec. 2018.

6. New Software and Platforms

6.1. io.datascience

Input Output Data Science

KEYWORDS: Open data - Semantic Web - FAIR (Findable, Accessible, Interoperable, and Reusable)

FUNCTIONAL DESCRIPTION: io.datascience (Input Output Data Science) is the instance of the Linked Wiki platform developed specifically in Paris-Saclay University as part of its Center for Data Science.

The goal of io.datascience: to facilitate the sharing and use of scientific data. The technological concept of io.datascience: the exploitation of semantic web advances, and in particular wiki technologies.

(Findable, Accessible, Interoperable, and Reusable) (Wilkinson, M., and The FAIR Guiding Principles for Scientific Data Management and Stewardship, Nature Scientific Data 2016)

io.datascience is both a data sharing platform and a framework for further development. It realizes a practical implementation of FAIR (Findable, Accessible, Interoperable, and Reusable - Wilkinson, M., Nature Scientific Data 2016) principles through a user-centric approach.

- Partners: Border Cloud - Paris Saclay Center for Data Science - Université Paris-Sud
- Contact: Cécile Germain-Renaud
- Publications: Data acquisition for analytical platforms: Automating scientific workflows and building an open database platform for chemical anlysis metadata - A platform for scientific data sharing - TFT, Tests For Triplestores - Une autocomplétion générique de SPARQL dans un contexte multi-services - Certifying the interoperability of RDF database systems - Transforming Wikipedia into an Ontology-based Information Retrieval Search Engine for Local Experts using a Third-Party Taxonomy - The Grid Observatory 3.0 - Towards reproducible research and open collaborations using semantic technologies
- URL: https://io.datascience-paris-saclay.fr/

6.2. Codalab

KEYWORDS: Benchmarking - Competition
FUNCTIONAL DESCRIPTION: Challenges in machine learning and data science are competitions running over several weeks or months to resolve problems using provided datasets or simulated environments. Challenges can be thought of as crowdsourcing, benchmarking, and communication tools. They have been used for decades to test and compare competing solutions in machine learning in a fair and controlled way, to eliminate "inventor-evaluator" bias, and to stimulate the scientific community while promoting reproducible science. See our slide presentation.

As of december 2017 there are 145 public competitions on Codalab and over 10000 users. Some of the areas in which Codalab is used include Computer vision and medical image analysis, natural language processing, time series prediction, causality, and automatic machine learning. Codalab was selected for the million Euro challenge See.4C that was awarded a H2020 EU grant for its organization.

TAU is going to continue expanding Codalab to accommodate new needs. One of our current focus is to support use of challenges for teaching (i.e. include a grading system as part of Codalab) and support for hooking up data simulation engines in the backend of Codalab to enable Reinforcement Learning challenges and simulate interactions of machines with an environment. For the third year, we are using Codalab for student projects. M2 AIC students create mini data science challenges in teams of 6 students. L2 math and informatics students then solve them as part of their mini projects. We are collaborating with RPI (New York, USA) to use this platform as part of a curriculum of medical students. Our PhD. students are involved in co-organizing challenges to expose the research community at large with the topic of their PhD. This helps them formalizing a task with rigor and allows them to disseminate their research.

- Partner: Microsoft
- Contact: Isabelle Guyon
- URL: http://competitions.codalab.org

6.3. Cartolabe

KEYWORD: Information visualization
FUNCTIONAL DESCRIPTION: The goal of Cartolabe is to build a visual map representing the scientific activity of an institution/university/domain from published articles and reports. Using the HAL Database, Cartolabe provides the user with a map of the thematics, authors and articles . ML techniques are used for dimensionality reduction, cluster and topics identification, visualisation techniques are used for a scalable 2D representation of the results.
NEWS OF THE YEAR: Improvement of the graphical interface

- Partners: LRI - Laboratoire de Recherche en Informatique - CNRS
- Contact: Philippe Caillou
- URL: http://www.cartolabe.fr/

7. New Results

7.1. Toward Good AI

7.1.1. Causal Modeling

Participants: Philippe Caillou, Isabelle Guyon, Michèle Sebag;
Post-docs and PhDs: Olivier Goudet, Diviyan Kalainathan
Collaboration: David Lopez-Paz (Facebook).

The search for **causal models** relies on quite a few hardly testable assumptions, e.g. causal sufficiency [152]; it is a data hungry task as it has the identification of independent and conditionally independent pairs of variables at its core. A new approach investigated through the Cause-Effects Pairs (CEP) Challenge [107] formulates causality search as a supervised learning problem, considering the joint distributions of pairs of variables (e.g. (Age, Salary)) labelled with the proper causation relationship between both variables (e.g. Age "causes" Salary) and learning algorithms apt to learn from distributions have been proposed [109]. An edited book is in preparation [64].

In D. Kalainathan's PhD and O. Goudet's postdoc, the search for causal models has been tackled in the framework of generative networks [44], trained to minimize the Maximum Mean Discrepancy loss; the resulting Causal Generative Neural Network improves on the state of the art on the CEP Challenge. However, due to the shortage of real-world variable pairs for which the causation type is known, the CEP challenge has been enriched using artificial pairs (e.g. considering variations on pairs of entities involved in biological regulatory networks), biasing the causation training process. On-going studies investigate how the use of such artificial pairs (the so-called Mother Distribution) to train a causation model aimed at real pairs can be cast as a domain adaptation problem [97], [78].

An attempt to circumvent the need for a large dataset of variable pairs, sampled for the Mother Distribution, we proposed the Structural Agnostic Model approach [57]. Working directly on the observational data, this global approach implements a variant of the popular adversarial game [97] between a discriminator, attempting to distinguish actual samples from fake ones, obtained by generating each variable, given real values from all others. A sparsity L_1 penalty forces all generators to consider only a small subset of their inpu variables, yielding a sparse causal graph. SAM obtains state-of-the-art performances on synthetic data.

An innovative usage of causal models is for educational training in sensitive domains, such as medicine, along the following line. Given a causal generative model, artificial data can be generated using a marginal distribution of causes; such data will enable students to test their diagnosis inference (with no misleading spurious correlations in principle), while forbidding to reverse-engineer the artificial data and guess the original data. Some motivating applications for causal modeling are described in section 4.1.

7.1.2. Explainability

Participants: Isabelle Guyon, François Landes, Marc Schoenauer, Michèle Sebag.

Causal modeling is one particular method to tackle explainability, and TAU has been involved in other initiatives toward explainable AI systems. Following the LAP (Looking At People) challenges, Isabelle Guyon and co-organizers have edited a book [29] that presents a snapshot of explainable and interpretable models in the context of computer vision and machine learning. Along the same line, they propose an introduction and a complete survey of the state-of-the-art of the explainability and interpretability mechanisms in the context first impressions analysis [56].

The team is also involved in the proposal for the IPL HyAIAI (Hybrid Approaches for Interpretable AI), coordinated by the LACODAM team (Rennes) dedicated to the design of hybrid approaches that combine state of the art numeric models (e.g., deep neural networks) with explainable symbolic models, in order to be able to integrate high level (domain) constraints in ML models, to give model designers information on ill-performing parts of the model, and to provide understandable explanations on its results.

Finally, a completely original approach to DNN explainability might arise from the study of structural glasses (7.2.3), with a parallel to CNNs with rotational invariances, that could become an excellent non-trivial example for developing explainability protocols.

7.1.3. Experimental Validation of the Autonomous Vehicle

Participants: Guillaume Charpiat, Marc Schoenauer; **PhD and Engineers**: Marc Nabhan, Nizham Makhoud, Raphaël Jaiswal
Collaboration: Hiba Hage, Philippe Reynaud, and Yves Tourbier (Renault)

As said (Section 3.1.2, TAU is considering two directions of research related to the certification of MLs. The first direction, toward experimental validation, focuses on the coverage of the datasets (more particularly here, used to train an autonomous vehicle controller), and is the subject of this section, while the second one, related to formal approaches, has just started with the beginning of Julien Girard's PhD and has not yet lead to results.

Statistical guarantees (e.g., less than 10^{-8} failure per hour of operation) are obtained by empirical tests, involving millions of kilometers of driving in all possible road, weather and traffic conditions as well as intensive simulations, the only way to full control of the driving conditions. The validation process thus involves 3 steps: i) making sure that all parts of the space of possible scenarios are covered by experiments/tests with sufficiently fine grain; ii) identify failures zones in the space of scenarios; iii) fix the controller flaws that resulted in these failures.

TAU is collaborating with Renault on steps i) (topic of a one-year POC) and ii) (Marc Nabhan's CIFRE PhD). In both cases, the current target scenario is the insertion of a car on a motorway, the "drosophila" of autonomous car scenarios.

Note that another approach toward experimental robustness is investigated in Nizam Makdoud's PhD (CIFRE Thalès), started in March 2018, where Reinforcement Learning is used to find ways to fool some security system.

Clustering of scenarios A first one-year Proof of Concept (ending Oct. 2018) has demonstrated the feasibility and the usefulness of scenario clustering, assuming the availability of data describing the scenarios, i.e., the trajectories of all vehicles involved. Publicly available datasets (e.g., NGSIM were used in a first step. The difficulties met are the following. Firstly, trajectories are varying-length time series, requiring the use of recurrent NNs or LSTMs. Secondly, a scenario is invariant under permutations of the different vehicles involved; neural architectures are taking inspiration from *social LSTMs* [67]. Lastly, most recorded real-world scenarios are uninteresting (all vehicles drive on in their lanes).

The results of this POC have been duly delivered to Renault, but will remain internal at this point. The follow-up collaboration will explore metrics (in the latent space, or learned via Siamese networks), to complete the clustering in a semi-supervised setting (exploiting human feedback to select "typical" scenarios).

Detection of controller flaws Marc Nabhan's PhD (CIFRE Renault) is concerned with the identification of the conditions of failures of the autonomous car controller. Only simulations are considered here, with one scenario being defined as a parameter setting of the in-house simulator SCANeR. The goal is the detection of as many failures as possible, running as few simulations as possible.

A key difficulty, beside that of getting actual data, is the very low probability of failure.On-going work builds upon TAU expertise in active learning using Monte-Carlo Tree Search [140] and evolutionary optimization, in particular taking inspiration from Novelty Search [121] to focus the exploration on unexplored regions of the scenario space, as well as portfolio optimization and instance-based algorithm selection (see Section 3.3.1).

7.2. Learning to Learn

*7.2.1. Auto-**

Participants: Guillaume Charpiat, Isabelle Guyon, Marc Schoenauer, Michèle Sebag
PhDs: Léonard Blier, François Gonard, Zhengying Liu, Herilalaina Rakotoarison, Lisheng Sun, Pierre Wolinski
Collaboration: Vincent Renault (SME Artelys); Olivier Bousquet (Google Zurich), Yann Ollivier (Facebook)

TAU is an active player in the Auto-☆ field, having organized the sixth COSEAL workshop in Paris in September 2018. Furthermore, Auto-☆ studies at TAU investigate several directions.

As discussed in Section 3.3, the most widely used approach is based on meta-features describing datasets, and builds upon past work in the team, such as Nacim Belkhir's PhD defended in 2017 [71], who won a GECCO competition in 2018 (Section 5.1.1), and François Gonard's PhD [11], defended in May 2018: an empirical performance model is built from the meta-features, and used to choose the best algorithm and its parameter configuration for unknown datasets. One key difficulty is to design useful meta-features: taking inspiration from equivariant learning [136] and learning from distributions [124], on-going work aims to learn such meta-features, based on the OpenML archive [153]. This extensive archive reports on the test predictive accuracy obtained by a few hundred algorithm configurations over a few thousand datasets.

Also mentioned in Section 3.3, another popular approach for algorithm selection is collaborative filtering. Active learning was used on top of the CofiRank algorithm for matrix factorization [156], improving the results and the time to solution of the recommendation algorithm [62].

An original approach to Auto-☆, explored in Herilalaina Rakotoarison's PhD, extends and adapts Monte-Carlo Tree Search to explore the structured space of pre-processing + learning algorithm configurations, and gradually determine the best pipeline [40]; the resulting algorithm yields promising results comparatively to AutoSklearn. A difficulty consists in managing the exploration together with the resource allocation (considering subsampled datasets and/or limited computational resources in the early MCTS stages, akin [91]).

Most real-world domains evolve with time, and an important issue in real-world applications is that of life-long learning, as static models can rapidly become obsolete. An extension of AutoSklearn was proposed, part of Lisheng Sun's PhD, that detects concept drifts and corrects the current model accordingly [38].

Two on-going works focus on the specific adjustment of hyper-parameters for neural nets, deriving rules for the network architecture (Pierre Wolinski's PhD), or (Leonard Blier's PhD) attaching fixed learning rates to each neuron and calibrating the learning rate distribution in such a way that neurons are sequentially active, learning in an optimally agile manner during a given learning phase, and being stable in later phases.

A last direction of investigation concerns the design of challenges, that contribute to the collective advance of research in the Auto-☆ direction. The team has been very active in the series of AutoML challenges [42], and continuously contributes to the organization of new challenges (Section 7.6).

7.2.2. Deep Learning: Practical Theoretical Insights

Participants: Guillaume Charpiat, Marc Schoenauer, Michèle Sebag
PhDs: Léonard Blier, Corentin Tallec
Collaboration: Yann Ollivier (Facebook AI Research, Paris), the Altschuler and Wu lab. (UCSF, USA)

Even though a full mathematical understanding of deep learning is not available today, theoretical insights from information theory or from dynamical systems can bring significant improvements to practical deep learning algorithms or offer strong explanations for the success of some architectures compared to others.

In [32] we fully derive the LSTM structure from first axiomatic principles, using an axiom of *robustness to temporal deformation (warpings) in the data*. The LSTM architecture, introduced in the 90's, has become the currently dominant architecture for modeling temporal sequences (such as text) in deep learning. But the LSTM architecture itself is quite complex and appears very much ad hoc at first sight. We prove that LSTMs necessarily arise if one wants the model to be able to handle time warpings in the data (such as arbitrary accelerations or decelerations in the signal). In fact, LSTM-like structures are the only way to provide robustness to such deformations: their complex equations can be derived axiomatically.

In [28] (long oral presentation at ICML) we tackle the problem of mode loss in generative models via information theory. The problem is to find generative models to produce more samples similar to samples in a dataset (eg, realistic images). The standard GAN approach is to couple a generative network and an adversary network whose job is to tell the differences between generated and genuine images. This suffers from mode loss: the generator focuses on doing some images well, rather than covering a full variety of images. Instead we propose to have the discriminator predict the proportion of true and fake images in a set of images, via an information theory criterion. This makes the discriminator work at the level of the *overall distribution* of images from the generator rather than individual images. By working on sets of images, the discriminator can detect statistical imbalances between different types of images created by the generator, thus reducing mode loss. An adapted architecture is derived for this, provably able to detect (in principle) all permutation-invariant statistics in a set of images.

In [43] we tackle the problem of recurrent network training via the theory of dynamical systems. Recurrent networks deal with temporal data sequences exhibiting temporal dependencies. Then backpropagation becomes backpropagation through time: for every new data point, training must rewind the network's computations backward in time on all past data to update the model parameters. This is unrealistic in any real-time application where the data arrive online. Two years ago we presented a fully online solution avoiding this "time rewind" step, based on real-time, noisy but unbiased approximations of model gradients. Our previous solution was mathematically well motivated but extremely complex to implement for standard models such as LSTMs. We now have a simpler variant which can be implemented easily in a black-box fashion on top of any recurrent model, and which is just as well-justified mathematically. The price to pay is more variance. In the long run, this could quite extend the applicability range of recurrent model to real-time situations.

In [31][0], we introduce a multi-domain adversarial learning algorithm in the semi-supervised setting. We extend the single source H-divergence theory for domain adaptation to the case of multiple domains, and obtain bounds on the average- and worst-domain risk in multi-domain learning. This leads to a new loss to accommodate semi-supervised multi-domain learning and domain adaptation. We obtain state-of-the-art results on two standard image benchmarks, and propose as a new benchmark a novel bioimage dataset, CELL, in the domain of automated microscopy data, where cultured cells are imaged after being exposed to known and unknown chemical perturbations, and in which each dataset displays significant experimental bias.

7.2.3. Analyzing and Learning Complex Systems

Participants: Cyril Furtlehner, Aurélien Decelle, François Landes
PhDs: Giancarlo Fissore
Collaboration: Jacopo Rocchi (LPTMS Paris Sud), the Simons team: Rahul Chako (post-doc), Andrea Liu (UPenn), David Reichman (Columbia), Giulio Biroli (ENS), Olivier Dauchot (ESPCI).
The information content of a trained restricted Boltzmann machine (RBM) for instance can be analyzed by comparing the singular values/vectors of its weight matrix, referred to as data modes, to that of a random RBM (typically following a Marchenko-Pastur distribution) [83]. The general strategy here is to replace the analysis of the learning process of a single instance by that of a well chosen statistical ensemble of models. In G. Fissore's PhD, the learning trajectory of an RBM is shown to start with a linear phase recovering the dominant modes of the data, followed by a non-linear regime where the interaction among the modes is characterized [15]. While the mean-field analysis conducted in closed form requires simplifying assumptions, it suggests some simple heuristics to speed up the convergence and to simplify the models. Ongoing works concern extensions of these considerations to settings with missing input on the practical side and to the analysis of exactly solvable RBM - i.e. non-linear RBM for which the contrastive divergence can be computed in closed forms - on the theoretical side. Additionnally, we are collaborating with J. Rocchi, working at the LPTMS (Univ. Paris Sud), to investigate the landscape of RBMs learned from different initial conditions and to characterized it as a function of the number of parameters (hidden nodes) of the system.

A long standing application of our aforementioned mean-field inference methods based on probabilistic modelling concerns road traffic forecasting. In [49] we wrap up some of the techniques developed in these past works and perform, thanks to PTV-SISTeMA comprehensive experimental tests on various real world Urban traffic dataset in order to illustrate in various conditions the effectiveness of our method. As a by-product we show to some extent how to disentangle the model bias from errors caused by corrupted data and shed some light on the nature of the data themselves.

[0] to be presented at ICLR 2019

An emerging research topic, that we started to investigate thanks to exchanges with Lenka Zdeborova's group [96], is to revisit the Information Bottleneck framework [151] and analyze on non-toy NNs the gradual distillation of the mutual information (MI) along the NN layers, minimizing the MI with the input while preserving the MI with the sought output (the labels). More generally, information theory concepts could also be used to analyze the behavior of the network, for instance to detect adversarial attacks through unusual neural activity mapping.

As mentioned earlier, the use of ML to address fundamental physics problems is quickly growing. One example is the domain of glasses (how the structure of glasses is related to their dynamics), which is one of the major problems in modern theoretical physics. The idea is to let ML models automatically find the hidden structures (features) that control the flowing or non-flowing state of matter, discriminating liquid from solid states. These models could then help identifying "computational order parameters", that would advance the understanding of physical phenomena, on the one hand, and support the development of more complex models, on the other hand. Furthermore, this problem is new to the ML community and could provide an original non-trivial example for engineering, testing and benchmarking explainability protocols.

7.3. Computational Social Sciences

Computational Social Sciences (CSS) is making significant progress in the study of social and economic phenomena thank to the combination of social science theories and new insight from data science. But while the simultaneous advent of massive data and massive computational power has opened exciting new avenues, it has also raised new questions and challenges.

Almost ten years after the first enthusiasms for "big data" in social science, P. Tubaro has undertaken a reflective effort to look back at progress made so far and at directions for the near future. She edited a special issue of *Revue Française de Sociologie* on the effects of data both on society itself and on the scientific disciplines that engage with it [46], of which she co-authored the introduction [13].

Meanwhile, four data-based studies are being conducted in TAU, about labor (hiring, working on Internet, quality of life and economic performance), about nutrition (health, food, and socio-demographic issues), around Cartolabe, a platform for scientific information system and visual querying and around GAMA, a multi-agent based simulation platform.

7.3.1. Labor Studies

Participants: Philippe Caillou, Isabelle Guyon, Michèle Sebag, Paola Tubaro
Post-docs; PhDs: Olivier Goudet; François Gonard, Diviyan Kalainathan, Thomas Schmitt
Collaboration: Jean-Pierre Nadal (EHESS); Marco Cuturi, Bruno Crépon (ENSAE); Antonio Casilli (Telecom); Thierry Weil (Mines); Jean-Luc Bazet (RITM)

A first area of activity of TAU in Computational Social Sciences is the study of labor, from the functioning of the job market, to the rise of new, atypical forms of work in the networked society of internet platforms, and the quality of life at work.
Job markets Our first study in the domain of job markets (Th. Schmitt's and F. Gonard's PhDs [12], [11]) tackled the matching of job ads and CVs. This study, funded by the Lidex *Institut de la Société Numérique* (ISN) at Univ. Paris-Saclay, was conducted in collaboration with EHESS, on data provided by the hiring Web agency Qapa (for blue-collars and temporary jobs) and by Association Bernard Gregory (for scientists in industry). Among other difficulties, this study revealed that for both qualified and unqualified job sectors, job seekers and recruiters do not speak the same language [143]. This first study will be continued and extended along two directions: counterfactual analysis (*What would be my options if I had this additional skill ?*DATAIA project Vadore, coll. ENSAE and Pôle Emploi), and the recommendation of vocational training (BPI-PIA contract JobAgile, coll. EHESS and Qapa). Both projects start end 2018.

The platform economy and digital labor Another topic concerns the digital economy and the transformations of labor that accompany the current developments of AI. P. Tubaro has researched the so-called "sharing economy" and ideals of social change associated to the economic model of the platform [33]. However, the platform economy is also disrupting traditional industries. CNRS's MITI office has funded a research on the effects of online services for the restaurant sector (such as La Fourchette, Trip Advisor, Yelp) on working conditions and quality of service. This project involves P. Tubaro, P. Caillou and partners at Telecom ParisTech and Paris Dauphine University.

Ongoing research is exploring online platform labor and its linkages to the development of AI. In collaboration with A.A. Casilli (Telecom ParisTech), P. Tubaro has received funding to conduct research on this topic from the Union Force Ouvrière (OPLa project), from France Stratégie (a Prime Minister's service), and from MSH Paris-Saclay (DiPLab project). A recent grant from DARES (French Ministry of Labor) will enable exploring labor changes in B2B platforms (with O. Chagny of IRES, a unions-funded think-tank).

Quality of life at work. A study, funded by ISN, examined the relationship between the quality of life at work (QLW), and the economic performance of companies [113]. The management and economics literature has already established a correlation between QLW and economic performance [76]. The question that we are currently addressing regards the direction of causality: do profitable companies pay more attention to the QLW ? Or do companies paying attention to QLW tend to be more profitable ? This project (coll. RITM Univ. Paris-Sud, SES Telecom ParisTech, Ecole des Mines, La Fabrique de l'Industrie) combines data at the individual level (DARES, Ministère du Travail) and at the company level (Secafi); cutting-edge causality algorithms are applied to address the question, and handle confounder variables such as the sector of activity.

7.3.2. Health, food, and socio-demographic issues

Participants: Philippe Caillou, Michèle Sebag, Paola Tubaro
Post-docs; PhDs: Nayat Sanchez-Pi
Collaboration: Louis-Georges Soler, Olivier Allais (INRA)

Another area of activity concerns the relationships between eating practices, socio-demographic features and health.

The Nutriperso project (IRS Univ. Paris-Saclay, coll. INRA, CEA, CNRS, INSERM, Telecom ParisTech and Univ. Paris-Sud) aims to: i) determine the impact of food items on health (e.g., related to T2 diabetes); ii) identify alternative food items, admissible in terms of taste and budget, and better in terms of health; iii) emit personalized food recommendations (noting that general recommendations such as *Eat 5 fruit and vegetable per day* are hardly effective on the targeted populations. Based on the Kantar database, reporting the food habits of 20,000 households over 20 years, our challenge is to analyze the food purchases at an unprecedented fine-grained scale (at the barcode level), and to investigate the relationship between diets, socio-demographic features, and body mass index (BMI). The challenge also regards the direction of causality; while some diets are strongly correlated to high BMI, the question is to determine whether, e.g., sugar-free sodas are a cause, or a consequence of obesity, or both.

Previous research in this area included the study of eating disorders and their relationship to people's social network and usages of technology [18].

7.3.3. Scientific Information System and Visual Querying

Participants: Philippe Caillou, Michèle Sebag
Engineer: Anne-Catherine Letournel, Jonas Renault
Collaboration: Jean-Daniel Fekete (AVIZ, Inria Saclay)

A third area of activity concerns the 2D visualisation and querying of the scientific expertise in an institute/university, based on their scientific production, given as a set of articles (authors, title, abstract). The Cartolabe project started as an Inria ADT (coll. TAO and AVIZ, 2015-2017). It received a grant from CNRS (coll. TAU, AVIZ and HCC-LRI, 2018-2019). Further extension proposals, in collaboration with the department of bibliometry from Univ. Paris-Saclay, are under submission at the time of writing.

This project was initially devised as an open-source platform, aimed to answer burning questions, as the growth of academic organization prevents anyone from having a precise knowledge of who does what in the organization: Who is expert in a topic (described as a bag of words)? How are topics related? What are the rising topics? (see also Section 6.3)

Its development and the interaction with the beta-user scientists using it, increasingly raises new questions at the crossroad of human-centered computing, data visualization and machine learning: How to deal with poly-thematic researchers? How to take advantage of the fact that researchers have ideas about their relevant scientific neighborhood, and learn person-dependent metric?

7.3.4. Multi-Agent based simulation framework for social science

Participants: Philippe Caillou
Collaboration: Patrick Taillandier (INRA), Alexis Drogoul and Nicolas Marilleau (IRD), Arnaud Grignard (MediaLab, MIT), Benoit Gaudou (Université Toulouse 1)

Since 2008, P. Caillou contributes to the development of the GAMA platform, a multi-agent based simulation framework. Its evolution is driven by the research projects using it, which makes it very well suited for social sciences studies and simulations.

The 1.8 version of the platform[20] brings new capabilities required for social science research, such as High Performance Computing to explore the simulation, Co-Modeling to link projects, advanced agent architectures to model complex behaviors and advanced visualization to display nice 3D representations for exploration and presentations.

7.4. Energy Management

7.4.1. Power Grids Daily Management

Participants: Isabelle Guyon, Marc Schoenauer
PhDs: Benjamin Donnot, Balthazar Donon, Herilalaina Rakotoarison
Collaboration: Antoine Marot, Patrick Panciatici (RTE), Olivier Teytaud (Facebook)

In the context of the Power Grid safety (Benjamin Donnot's CIFRE PhD with RTE, to be defended in February 2019), the goal is to assess in real time the so-called "(n-1)" safety (see Section 4.2) of possible recovery actions after some problem occurred somewhere on the grid. However, the simulator that allows to compute the power flows in the whole network is far too slow to simulate in real time all n-1 possible failures. A simplified simulator is also available, but its accuracy is too poor to give any good result. Deep surrogate models can be trained off-line, based on the results of the slow simulator, with high0enough accuracy, but training as many models as possible failures (i.e., n-1), obviously doesn't scale up: the topology of the grid must be an input of the learned model, allowing to instantly compute the power flows at least for grid configurations close to the usual running state of the grid. A standard approach is the one-hot encoding of the topology, where n additional boolean inputs are added to the neural network, encoding the presence or absence of each line. An original "guided dropout" approach was proposed [24], in which the topology directly acts on the connections of the deep network: a missing line suppresses some connections. However, whereas the standard dropout method disconnect random connections for every batch, in order to improve the generalization capacity of the network, the "guided dropout" method removes some connections based on the actual topology of the network. THis approach is experimentaly validated against the above-mentionned approaches on small subsets of the French grid (up to 308 lines). Interestingly, and rather suprisingly, even though only examples with a single disconnected line are used in the training set, the learned model is able of some additive generalization, and predictions are also accurate enough in the case 2 lines are disconnected. The guided dropout approach was later robustified [23] by learning to rapidly rank higher order contingencies including all pairs of disconnected lines, in order to prioritize the cases where the slow simulator is run: Another neural network is trained to rank all (n-1) and (n-2) contingencies in decreasing order of presumed severity.

7.4.2. Local Grids Optimization, and the Modeling of Worst-case Scenarios

Participants: Isabelle Guyon, Marc Schoenauer, Michèle Sebag
PhDs: Victor Berger, Herilalaina Rakotoarison; **Post-doc**: Berna Batu
Collaboration: Vincent Renaut (Artelys)

One of the goals of the ADEME Next project, in collaboration with SME Artelys (see also Section 4.2), is the sizing and capacity design of regional power grids. Though smaller than the national grid, regional and urban grids nevertheless raise scaling issues, in particular because many more fine-grained information must be taken into account for their design and predictive growth.

Provided accurate predictions of consumption (see below), off-the-shelf graph optimization algorithms can be used. Berna Batu is gathering different approaches, while Herilalaina Rakotoarison's PhD is concerned with the automatic tuning of their parameters (see Section 7.2.1, and his original approach, at the moment applied to standard benchmarks [40], as well as to Artelys' home optimizer at large Knitro, and compared to the state-of-the-art in parameter tuning (confidential deliverable).

In order to get accurate consumption predictions, V. Berger's PhD tackles the identification of the peak of energy consumption, defined as the level of consumption that is reached during at least a given duration with a given probability, depending on consumers (profiles and contracts) and weather conditions. The peak identification problem is currently tackled using Monte-Carlo simulations based on consumer profile- and weather-dependent individual models, at a high computational cost. The challenge is to exploit individual models to train a generative model, aimed to sampling the collective consumption distribution in the quantiles with highest peak consumption.

7.5. Data-driven Numerical Modelling

7.5.1. High Energy Physics

Participants: Cécile Germain, Isabelle Guyon
PhD: Victor Estrade, Adrian Pol
Collaboration: D. Rousseau (LAL), M. Pierini (CERN)

The role and limits of simulation in discovery is the subject of V. Estrade's PhD, specifically uncertainty quantification and calibration, that is how to handle the systematic errors, arising from the differences ("known unknowns") between simulation and reality, coming from uncertainty in the so-called nuisance parameters. In the specific context of HEP analysis, where relatively numerous labelled data are available, the problem is at the crosspoint of domain adaptation and representation learning. We have investigated how to directly enforce the invariance w.r.t. the nuisance in the sought embedding through the learning criterion (tangent back-propagation) or an adversarial approach (pivotal representation). The results [25] contrast the superior performance of incorporating a priori knowledge on a well separated classes problem (MNIST data) with a real case setting in HEP, in relation with the Higgs Boson Machine Learning challenge [66]. More indirect approaches based on either incorporating variance reduction for the parameter of interest or constraining the representation in a variational auto-encoder farmework are currently considered.

Anomaly detection is the subject of A. Pol PhD. Reliable data quality monitoring is a key asset in delivering collision data suitable for physics analysis in any modern large-scale high energy physics experiment. [60] focuses on supervised and semi-supervised methods addressing the identification of anomalies in the data collected by the CMS muon detectors. The combination of DNN classifiers capable of detecting the known anomalous behaviors, and convolutional autoencoders addressing unforeseen failure modes has shown unprecedented efficiency, compared either to production solution or classical anomaly detection (one-class or I-Forest). The result has been included in the production suite of the CMS experiment at CERN.

The highly visible TrackML challenge is described in section 7.6.

7.5.2. Remote Sensing Imagery

Participants: Guillaume Charpiat

Collaboration: Yuliya Tarabalka, Armand Zampieri, Nicolas Girard, Pierre Alliez (Titane team, Inria Sophia-Antipolis)

The analysis of satellite or aerial images has been a long-time ongoing topic of research, but the remote sensing community moved only very recently to a principled vision of the tasks in a machine learning perspective, with sufficiently large benchmarks for validation. The main topics are the segmentation of (possibly multispectral) remote sensing images into objects of interests, such as buildings, roads, forests, etc., and the detection of changes between two images of the same place taken at different moments. The main differences with classical computer vision is that images are large (covering whole countries, typically cut into 5000×5000 pixels tiles), containing many small, potentially similar objects (and not one big object per image), that every pixel needs to be annotated (w.r.t. assigning a single label to a full image), and that the ground truth is often not reliable (spatially mis-registered, missing new constructions).

This year, deep learning techniques took over classical approaches in most labs, adapting neural network architectures to the specifics of the tasks. This is due notably to the creation of several large scale benchmarks (including one by us [127] and, soon after, larger ones by GAFAM). A still ongoing issue is the ability to generalize across datasets (as urban and rural areas look different in different parts of the world, or even within the same country, e.g. roof types in France).

The task of segmenting satellite images comes together with the one of their registration with cadastral maps. Indeed, the ground truth in remote sensing benchmarks (cadastral maps) is often imperfect, due to spurious deformations. We tackle this issue by *learning* how to register images of different modalities (RGB pictures vs. binary cadastral maps). If one tries to predict, given an RGB photography and an associated cadastral map, the deformation that warps one onto the other, by outputting a 2D vector field indicating the predicted displacement of each pixel (which can be as large as ± 32 px), then the problem considered is too hard (32×32 possibilities for each pixel 2D displacement vector). Instead, we simplifiy the problem by decomposing it in a cascade of increasing resolutions. The idea is that if one zooms out by a factor 32, while knowing that the maximum possible displacement is of magnitude 32 px, then at this low resolution one has to move pixels by at most 1 pixel. Learning the task at this low resolution is thus easy. When it is done, if we zoom in by a factor 2, thus reaching a resolution lower than the original one by a factor 16, then the maximum displacement is again of 1 pixel (since larger displacements have been dealt with at the previous scale). And so on. In the end, we train a multi-scale chain of neural networks (double U-nets) [34], and later combine it with a segmentation task [27] in order to benefit from multi-task training, known to improve results.

7.5.3. Space Weather Forecasting

Participants: Cyril Furtlehner, Michèle Sebag
PhD: Mandar Chandorkar
Collaboration: Enrico Camporeale (CWI)

Space Weather is broadly defined as the study of the relationships between the variable conditions on the Sun and the space environment surrounding Earth. Aside from its scientific interest from the point of view of fundamental space physics phenomena, Space Weather plays an increasingly important role on our technology-dependent society. In particular, it focuses on events that can affect the performance and reliability of space-borne and ground-based technological systems, such as satellite and electric networks that can be damaged by an enhanced flux of energetic particles interacting with electronic circuits.[0]

Since 2016, in the context of the Inria-CWI partnership, a collaboration between TAU and the Multiscale Dynamics Group of CWI aims to long-term Space Weather forecasting. The project is extremely timely, as the huge amount of (freely available) space missions data has not yet been systematically exploited in the current computational methods for space weather. Specifically, the goal is to take advantage of the data produced everyday by satellites surveying the sun and the magnetosphere, and more particularly to relate solar images and the quantities (e.g., electron flux, proton flux, solar wind speed) measured on the L1 libration

[0] After a recent survey conducted by the insurance company Lloyd's, an extreme Space Weather event could produce up to $2.6 trillion in financial damage.

point between the Earth and the Sun (about 1,500,000 km and 1 hour time forward of Earth). The project is very ambitious: the accurate prediction of e.g., geomagnetic storms, or solar wind speed from solar images, would represent a giant leap in the field. A challenge is to formulate such goals in terms of supervised learning problem, while the "labels" associated to solar images are recorded at L1 (thus with a varying and unknown time lag). In essence, while typical ML models aim to answer the question *What*, our goal here is to answer both questions *What* and *When*. Concerning the prediction of solar wind impacting earth magnetosphere from solar images, we encountered an interesting sub-problem related to the non deterministic travel time of a solar eruption to earth's magnetosphere. We have formalized it as the joint regression task of predicting the magnitude of signals as well as the time delay with respect to their driving phenomena and provided a solution tested on synthetic data.

7.5.4. Genomic Data and Population Genetics

Participants: Guillaume Charpiat, Flora Jay
PhD: Théophile Sanchez
Collaboration: TIMC-IMAG (Grenoble), Estonian Biocentre (Institute of Genomics, Tartu, Estonia)

Thanks to the constant improvement of DNA sequencing technology, large quantities of genetic data should greatly enhance our knowledge about evolution and in particular the past history of a population. This history can be reconstructed over the past thousands of years, by inference from present-day individuals: by comparing their DNA, identifying shared genetic mutations or motifs, their frequency, and their correlations at different genomic scales. Still, the best way to extract information from large genomic data remains an open problem; currently, it mostly relies on drastic dimensionality reduction, considering a few well-studied population genetics features.

On-going work at TAU, around Théophile Sanchez' PhD, co-supervised by G. Charpiat and Flora Jay, aims at extracting information from genomic data using deep neural networks; the key difficulty is to build flexible problem-dependent architectures, supporting transfer learning and in particular handling data with variable size. In collaboration with the Bioinfo group at LRI, we designed new generic architectures, that take into account DNA specificities for the joint analysis of a group of individuals, including its variable data size aspects [141]. In the short-term these architectures can be used for demographic inference; the longer-term goal is to integrate them in various systems handling genetic data (e.g., epidemiological statistics) or other biological sequence data. In collaboration with the Estonian Biocentre (Tartu, Estonia), applications will consider thousands of sequenced human genomes, and expand our knowledge of the past human history. To this aim Burak Yelmen (PhD student at the Estonian Biocentre) will visit the lab from February to April 2019. Indeed, TAU expertise regarding the methodologies of exploiting missing and noisy data, and the resulting modeling biases, can contribute to enhance these novel population genetics methods, particularly so for methods heavily relying on simulated data (thus potentially suffering from the *reality gap*).

We also contributed to tess, a method for fast inference of population genetic structure, through a collaboration with TIMC-IMAG. This method analyses SNP data and estimates the admixture coefficients (that is, the probability that an individual belongs to different groups given the genetic data) via matrix factorization. The observed high dimensional genetic data are reduced automatically via the rank-k approximation of the matrix factorization and thereby highlight the latent structure of the data: the matrix factorization scores correspond to the admixture coefficients while the loadings give the genetic characteristics of each cluster. This method is faster than the hierarchical Bayesian models that we had previously developed and hence well suited for large NGS data. We participated in the tess3 R package, that implements this algorithm, facilitates the visualization of population genetic structure and the projection on maps [14]. We are currently adapting closely related algorithms to enable dimension reduction of temporal data with an application to paleogenomics.

7.5.5. Sampling molecular conformations

Participants: Guillaume Charpiat
PhD: Loris Felardos
Collaboration: Jérôme Hénin (IBPC), Bruno Raffin (InriAlpes)

Numerical simulations on massively parallel architectures, routinely used to study the dynamics of biomolecules at the atomic scale, produce large amounts of data representing the time trajectories of molecular configurations. The configuration space is high-dimensional (10,000+), hindering the use of standard data analytics approaches. The use of advanced data analytics to identify intrinsic configuration patterns could be transformative for the field.

The high-dimensional data produced by molecular simulations live on low-dimensional manifolds; the extraction of these manifolds will enable to drive detailed large-scale simulations further in the configuration space. Among the possible options are i) learning a parameterization of the local, low-dimensional manifold and performing a geometric extrapolation of the molecule trajectories; ii) learning a coarse description of the system and its dynamics, supporting a fast prediction of its evolution. In both cases, the states estimated from the time- or configuration-simplified models will be used for steering large scale simulations, thus accelerating the sampling of stable molecular conformations.

This task will be tackled by combining manifold learning (to find a relevant low-dimensional representation space) and reinforcement learning (for the efficient exploration of the space), taking inspiration from Graph Neural Networks [86]. On-going studies use Graph Auto-encoders to extract a meaningful representation of the conformation of molecules and to predict dynamics.

7.5.6. Storm trajectory prediction

Participants: Mo Yang, Guillaume Charpiat
Collaboration: Claire Monteleoni, Sophie Giffard-Roisin (LAL / Boulder University), Balazs Kegl (LAL)

Cyclones, hurricanes or typhoons all designate a rare and complex event characterized by strong winds surrounding a low pressure area. Their trajectory and intensity forecast, crucial for the protection of persons and goods, depends on many factors at different scales and altitudes. Additionally storms have been more numerous since the 1990s, leading to both more representative and more consistent error statistics.

Currently, track and intensity forecasts are provided by numerous guidance models. Dynamical models solve the physical equations governing motions in the atmosphere. While they can provide precise results, they are computationally demanding. Statistical models are based on historical relationships between storm behavior and other parameters [82]. Current national forecasts are typically driven by consensus methods able to combine different dynamical models.

Statistical models perform poorly compared to dynamical models, although they rely on steadily increasing data resources. ML methods have scarcely been considered, despite their successes in related forecasting problems [160]. A main difficulty is to exploit spatio-temporal patterns. Another difficulty is to select and merge data coming from heterogeneous sensors. For instance, temperature and pressure are real values on a 3D spatial grid, while sea surface temperature or land indication rely on a 2D grid, wind is a 2D vector field, while many indicators such as geographical location (ocean, hemisphere...) are just real values (not fields), and displacement history is a 1D vector (time). An underlying question regards the *innate* vs *acquired* issue, and how to best combine physical models with trained models. On-going studies, conducted in collaboration with S. Giffard-Roisin and C. Monteleoni (now Univ. Boulder), outperform the state-of-the-art in many cases [26], [36], [35].

7.5.7. Analyzing Brain Activity

Participants: Guillaume Charpiat
Collaboration: Hugo Richard, Bertrand Thirion (Parietal team, Inria Saclay / CEA)

With the goal of understanding brain functional architecture, the brain activity of ten subjects is recorded by an fMRI scanner, while they are watching movies (sequences of short pieces of real movies). The analysis of the ensuing complex stimulation streams proceeds by extracting relevant features from the stimuli and correlating the occurrence of these features with brain activity recorded simultaneously with the presentation of the stimuli. The analysis of video streams has been carried in [87] or [108] using a deep convolutional network trained for image classification. The question is then to build good descriptors of videos, possibly involving motion.

We consider a deep neural network trained for action recognition on the largest dataset available [115], and use its activations as descriptors of the input video. This provides deep representations of the watched movies, from an architecture that relies either on optical flow, or on image content, or both simultaneously. We then train a linear model to predict brain activity from these features. From the different layers of the deep neural networks, we build video representations that allow us to segregate (1) occipital and lateral areas of the visual cortex (reproducing the results of [108] and (2) foveal and peripheric areas of the visual cortex. We also introduce an effi- cient spatial compression scheme for deep video features that allows us to speed up the training of our predictive algorithm [41]. We show that our compression scheme outperforms PCA by a large margin.

7.6. Challenges

Participants: Cécile Germain, Isabelle Guyon, Michèle Sebag
PhD: Zhengying Liu, Lisheng Sun
Collaboration: D. Rousseau (LAL), Andre Elisseeff (Google Zurich), Jean-Roch Vilmant (CERN)

Following the highly successful ChaLearn **AutoML** Challenges (NIPS 2015 – ICML 2016 [106] – PKDD 2018 [45]), the **AutoDL** challenge [37], to be run in 2019, addresses the problem of tuning the hyperparameters of Deep Neural Networks, including the topology of the network itself. Co-sponsored by Google Zurich, it will require participants to upload their code on the Codalab platform.

In conjunction with AutoDL, we will organize a challenge in computer vision called **AutoCV**, to promote automatic machine learning for video processing, in collaboration with University of Barcelona. This will make use of the TAU GPU cluster.

Part of the HEP activities of the team, **TrackML** [30], [61] first phase was run and co-sponsored by Kaggle, until September 2018. The second phase is presently running on Codalab, and will end in March 2019. The challenge has been presented at WCCI [61] and NIPS [30]. I. Guyon and C. Germain are in the organizing committee, and M. Schoenauer is member of the Advisory Committee. The TAU team, in collaboration with CERN, has taken a leading role in stimulating both the ML and HEP communities to address the combinatorial complexity explosion created by the next generation of particle detectors.

Beyond the LAP (Looking At People) series of challenges (see details and references in Section 3.4), the domain of autonomous analysis of human behavior from multimodal information has recently gained momentum. We have been involved in two Special Issues dedicated to these topics, *The Computational Face*, in PAMI [17], and *Apparent Personality Analysis*, in IEEE Trans. on Affective Computing [16]. Two other challenges were organized at ICPR 2018, one about the information fusion task in the context of multi-modal image retrieval in social media, the other one regarding the inference of personality traits from written essays, including textual and handwritten information [29].

The **HADACA** project (EIT Health) aims to run a series of challenges to promote and encourage innovations in data analysis and personalized medicine. The data challenges will gather transdisciplinary instructors (re-searchers and professors), students, and health professionals (clinicians). The outcome of the data challenges should provide: i) analytical frameworks to bridge the gap between large dataset and personalized medicine in disease treatments and ii) innovative pedagogical methods to sensitize students to big data analysis in health. As a synergistic activity, TAU is also engaged in a collaboration with the Rensselaer Polytechnic Institute (RPI, New-York, USA) to use challenges in the classroom, as part of their health-informatics curriculum.

The **L2RPN** (Learning to Run a Power Network) project (coll. RTE) [39] addresses the difficult problem of using Reinforcement Learning to assist human operator in their daily tasks of maintaining the French Ultra-High Voltage grid safety while routing power without interruption. We are collaborating with O. Pietquin (Google Brain) to firm up the challenge protocol, largely inspired by AlphaGo and other RL challenges, like the NIPS 2017 “Learning to run” challenge.

It is important to introduce **challenges in ML teaching**. This has been done (and is on-going) in I. Guyon's Licence and Master courses: some assignments to Master students are to design small challenges, which are then given to Licence students in labs, and both types of students seem to love it. Along similar line, F. Landes proposed a challenge in the context of S. Mallat's course, at Collège de France.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

TAU will continue TAO policy about technology transfer, accepting any informal meeting following industrial requests for discussion (and we are happy to be too much solicited), and deciding about the follow-up based upon the originality, feasibility and possible impacts of the foreseen research directions, provided they fit our general canvas. This lead to the following 5 on-going CIFRE PhDs, with the corresponding side-contracts with the industrial supervisor, plus 3 other bilateral contracts. In particular, we now have a first "Affiliate" partner, the SME DMH, and hope to further develop in the future this form of transfer. Note that it can also sometimes lead to collaborative projects, as listed in the following sections.

- **CIFRE RTE** 2015-2018 (72 kEuros), with Réseau Transport d'Electricité, related to Benjamin Donnot's CIFRE PhD
  Coordinator: Olivier Teytaud (until May 2016), then Isabelle Guyon, and Antoine Marot (RTE)
  Participants: Benjamin Donnot, Marc Schoenauer
- **Myndblue**, 2017-2018 (1 an, 50kEuros) related to consulting activities with DMH (Digital for Mental Health)[0].
  Coordinator: Aurélien Decelle and Simon Moulieras (DMH)
  Participants: Michèle Sebag
- **Contrat LFI** 2017-2018 (30kEuros), with La Fabrique de l'Industrie, related to quality of life at work (Section 7.3.1).
  Coordinator: Michèle Sebag and Thierry Weil (La Fabrique de l'Industrie)
  Participants: Olivier Goudet, Diviyan Kalainathan
- **POC Renault** 2017-2018 (125 kEuros), *Clusterisation et optimisation de scenarii pour la validation des véhicules autonomes*
  Coordinator: Marc Schoenauer and Philippe Reynaud (Renault)
  Participants: Guillaume Charpiat, Raphaël Jaiswal (engineer), Marc Schoenauer
- **CIFRE Renault** 2017-2020 (45 kEuros), related to Marc Nabhan's CIFRE PhD *Sûreté de fonctionnement d'un véhicule autonome - évaluation des fausses détections au travers d'un profil de mission réduit*
  Coordinator: Marc Schoenauer and Hiba Hage (Renault)
  Participants: Marc Nabhan (PhD), Yves Tourbier (Renault)
- **OPLa** 2017-2018, Organizing Platform Labor (27k euros), funded by Force Ouvrière.
  Coordinator: A.A. Casilli (Telecom ParisTech)
  Participants: Paola Tubaro
- **DiPLab** 2017-2018, Digital Platform Labor (24k euros), funded by MSH Paris-Saclay.
  Coordinators: Paola Tubaro (avec A.A. Casilli, Telecom ParisTech)
- **CIFRE Thalès** 2018-2021 (45 kEuros), with Thales Teresis, related to Nizam Makdoud's CIFRE PhD
  Coordinator: Marc Schoenauer and Jérôme Kodjabatchian
  Participants: Nizam Makdoud
- **CIFRE RTE** 2018-2021 (72 kEuros), with Réseau Transport d'Electricité, related to Balthazar Donon's CIFRE PhD

[0]This "Affiliate" contract has been inspired by the affiliate program of Technion

Coordinator: Isabelle Guyon and Antoine Marot (RTE)
Participants: Balthazar Donon, Marc Schoenauer

- **CIFRE FAIR** 2018-2021 (45 kEuros), with Facebook AI Research, related to Leonard Blier's CIFRE PhD
  Coordinator: Marc Schoenauer and Yann Olliver (Facebook)
  Participatns: Guillaume Charpiat, Michèle Sebag, Léonard Blier
- **Google Zurich** 2018 (50kEuros), related to the **AutoDL** (see Section 3.4)
  Coordinator: Isabelle Guyon and Olivier Bousquet (Google)
  Participants: Zhengying Liu and Lisheng Sun
- **IFPEN** (Institut Français du Pétrole Energies Nouvelles) 2018-2022 (300 kEuros), to hire an Inria Starting Research Position (PhD + 4-6 years) to work in all topics mentioned in Section 3.2 relevant to IFPEN activity (see also Section 4.2).

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

- **ACTEUR** 2014-2018 (236kEuros). Cognitive agent development for urban simulations,
  Coordinator: P. Taillandier (IDEES, Univ Rouen)
  Participant: Philippe Caillou
- **EPITOME** 2017-2020 (225kEuros), *Efficient rePresentatIon TO structure large-scale satellite iMagEs* (Section 7.5.2).
  Coordinator: Yuliya Tarabalka (Titane team, Inria Sophia-Antipolis)
  Participant: Guillaume Charpiat

9.1.2. Others

- **E-LUCID** 2014-2018 (194 kEuros), anomaly detection in network packets.
  Coordinator: Thales Communications & Security S.A.S
  Participants: Marc Schoenauer, Cyril Furtlehner, Luis Marti (until 12/2017)
- **Nutriperso** 2017-2020, 87 kEuros. Personalized recommendations toward healthier eating practices (Section 7.3.2).
  U. Paris-Saclay IRS (*Initiative de Recherche Stratégique*)
  Partners: INRA (coordinator), INSERM, Agro Paristech, Mines Telecom
  Participants: Philippe Caillou, Flora Jay, Michèle Sebag, Paola Tubaro
- **PIA Adamme** 2015-2018 (258 kEuros) Machine Learning on a mass-memory architecture.
  Coordinator: Bruno Farcy (Bull SAS)
  Participants: Marc Schoenauer, Guillaume Charpiat, Cécile Germain-Renaud, Yasmina Bouzbiba, Etienne Brame
- **NEXT** 2017-2021 (675 kEuros). Simulation, calibration, and optimization of regional or urban power grids (Section 4.2).
  ADEME (Agence de l'Environnement et de la Maîtrise de l'Energie)
  Coordinator: ARTELYS
  Participants Isabelle Guyon, Marc Schoenauer, Michèle Sebag, Victor Berger (PhD), Herilalaina Rakotoarison (PhD), Berna Bakir Batu (Post-doc)

- **DATAIA Vadore** 2018-2020 (105 kEuros) VAlorizations of Data to imprOve matching in the laboR markEt, with CREST (ENSAE) and Pôle Emploi (Section 7.3.1).
  Coordinator: Michèle Sebag
  Participants: Philippe Caillou, Isabelle Guyon
- **PIA JobAgile** 2018-2021 (379 kEuros) *Evidence-based Recommandation pour l'Emploi et la Formation* (Section 7.3.1).
  Coordinator: Michèle Sebag and Stéphanie Delestre (Qapa)
  Participants: Philippe Caillou, Isabelle Guyon
- **HADACA** 2018-2019 (50 kEuros), within EIT Health, for the organization of challenges toward personalized medicine (Section 7.6).
  Coordinator: Magali Richard (Inria Grenoble)
  Participants: Isabelle Guyon
- **IPL HPC-BigData** 2018-2022 (100 kEuros) High Performance Computing and Big Data (Section 7.5.5)
  Coordinator: Bruno Raffin (Inria Grenoble)
  Participants: Guillaume Charpiat, Loris Felardos (PhD)

9.2. European Initiatives

9.2.1. Collaborations with Major European Organizations

MLSpaceWeather 2015-2019. Coupling physics-based simulations with Artificial Intelligence (Section 7.5.3).
Coordinator: CWI
Participants: Aurélien Decelle, Cyril Furtlehner, Michèle Sebag

9.3. International Initiatives

9.3.1. Inria International Labs

IIL CWI-Inria
Associate Team involved in the International Lab:

9.3.1.1. MDG-TAO

Title: Data-driven simulations for Space Weather predictions

International Partner (Institution - Laboratory - Researcher):

CWI (Netherlands) - Multiscale Dynamics Group - Enrico Camporeale

Start year: 2017

See also: http://pages.saclay.inria.fr/cyril.furtlehner/html/mdg-tao.html and Section 7.5.3.

We propose an innovative approach to Space Weather modeling: the synergetic use of state-of-the-art simulations with Machine Learning and Data Assimilation techniques, in order to adjust for errors due to non-modeled physical processes, and parameter uncertainties. We envision a truly multidisciplinary collaboration between experts in Computational Science and Data assimilation techniques on one side (CWI), and experts in Machine Learning and Data Mining on the other (Inria). Our research objective is to realistically tackle long-term Space Weather forecasting, which would represent a giant leap in the field. This proposal is extremely timely, since the huge amount of (freely available) space missions data has not yet been systematically exploited in the current computational methods for Space Weather. Thus, we believe that this work will result in cutting-edge results and will open further research topics in space Weather and Computational Plasma Physics.

9.3.2. Inria International Partners

9.3.2.1. Declared Inria International Partners

Isabelle Guyon partner of Google Zurich *Preparation of a competition AutoDL: Automatic Deep Learning*. See Section 7.6.

9.3.2.2. Informal International Partners

Marc Schoenauer partner of the ARC-DP (Australian Research Council Discovery Project) *Bio-inspired computing methods for dynamically changing environments*. Coordinator: University of Adelaide (Frank Neumann), 5 years from Nov. 2015, 400 k$-AUS. Visit to Adelaide: 2 weeks in Feb. 2017, 2 weeks planned in 2019.

Isabelle Guyon Partner of UC Berkeley *Fingerprint verification with deep siamese neural networks using ultratonic sensor data.* Co-advisor of a master student (Baiyu Chen). Partners: Alyosha Efros, Bernhard Boser.

Guillaume Charpiat partner of Boulder University *Hurricane trajectory prediction*. Co-advisor of a master student (Mo Yang). Partners: Sophie Giffard-Roisin, Claire Monteleoni. See Section 7.5.6.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

Guillaume Charpiat Workshop Statistics/Learning at Paris-Saclay 2018

Isabelle Guyon Competition co-chair, ECML 2019

Michele Sebag Séminaire Annuel Académie des Technologies, 2018.

In 2016, Isabelle Guyon was Program Chair of NIPS (Neural Information Processing Systems) – with an increase of more than 40% in the number of submissions, 96% in terms of reviewers, and over 100% in terms of attendees as compared to the previous year. Lessons learned from the review process are described in a JMLR paper [19].

10.1.1.2. Member of Organizing Committees

Isabelle Guyon Advisory committee BayLearn 2018; Co-organizer WCCI 2018 Special Session on Intelligent Power Systems; Co-organizer WCCI 2018 Special Session on Machine Learning and Deep Learning Methods applied to Vision and Robotics (MLDLMVR); Co-organizer ECCV 2018 workshop Chalearn Looking at People: Inpainting and Denoising in the Deep Learning Age; Co-organizer 2018 Multimedia Information Processing for Personality & Social Networks Analysis Workshop at ICPR; Co-organizer NeurIPS 2018 workshop on Challenges in Machine Learning;

Marc Schoenauer Steering Committee, Parallel Problem Solving from Nature (PPSN); Steering Committee, Learning and Intelligent OptimizatioN (LION); organizer (with Herilalaina Rokotoarison), Workshop COSEAL, Paris, Sept. 17-18, 2018.

Michele Sebag President of Steering Committee, Eur. Conf. on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD).

10.1.2. Scientific Events Selection

10.1.2.1. Member of Conference Program Committees

All TAU members are members of the Program Committees of the main conferences in their respective fields of expertise.

10.1.2.2. Reviewer

All TAU member review papers for the most prestigious conferences in their respective fields of expertise.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

Isabelle Guyon Action editor, *Journal of Machine Learning Research* (JMLR); series editor, *Springer series Challenges in Machine Learning* (CiML).

Marc Schoenauer Advisory Board, *Evolutionary Computation Journal*, MIT Press, and *Genetic Programming and Evolutionary Machines*, Springer Verlag; Action editor, *Journal of Machine Learning Research*(JMLR).

Michèle Sebag Editorial Board, Machine Learning, Springer Verlag.

Paola Tubaro Associate Editorial Board, *Sociology*, Sage; Member of Editorial Board, *Revue Française de Sociologie*, Presses de Sciences Po.

10.1.3.2. Reviewer - Reviewing Activities

All members of the team reviewed numerous articles for the most prestigious journals in their respective fields of expertise.

10.1.4. Invited Talks

Guillaume Charpiat *Introduction aux réseaux de neurones*, Séminaire Parisien de Mathématiques Appliquées à l'Imagerie, Paris, 3 May. 2018; *Recalage et mise à jour d'images à l'aide de réseaux de neurones*, Journée Extraction d'attributs et apprentissage pour l'analyse des images de télédétection, GDR ISIS, Paris, 18 Oct. 2018.

Flora Jay *Deep Learning Methods for Population Genetics: Inferring Changes Population Size*, Journée Inférences évolutives, GDR GE et AEIM, Paris, 16 Mai 2018; *Inferring past history from genetic data using ABC and Deep Learning approaches*, INRA GenPhySE Seminar, Toulouse, 12 Dec. 2018.

Isabelle Guyon *Codalab: crowdsourcing DataIA* DATAIA Institute Kick-Off (Data Science, Intelligence & Society), 15 Feb. 2018; *Contests of contests*, La Recherche scientifique « hors murs » au 21e siècle, Colloque de l'Académie des sciences, Fondation Del Duca, 29 Nov. 2018; *Evaluating causation coefficients*, NeurIPS 2018 workshop on causal learning, 7 Dec, 2018.

Michèle Sebag *Qualité de la vie et santé économique, étude causale*: Colloque de Cerisy (Sept. 18); séminaire Université Québec à Montréal (Dec. 18); *Causal Modeling*: KAUST Conference on Computational and Statistical Interface to Big Data (Mar. 18); Leiden wshop on Space Weather (2018); *MonteCarlo Tree Search for Algorithm Selection and Calibration*: Dagstuhl Seminar (Sept. 18); NIPS Wshop on AutoML (Dec. 18); *Ingénieures et Scientifiques*, ENPC (Juin 18). Exposés Journées Cabourg, Mar. 18; Rennes, Sep. 18; Toulouse, Oct. 18.

Marc Schoenauer *The Villani mission on Artificial Intelligence*, CISCO Heaadquarters, 9 Apr. 2018; *Le rapport Villani*, journée AFIA, 12 Apr. 2018; *Intelligence Artificielle : le rapport Villani*, Journée de la Recherche de l'Université de Brest, 25 May 2018; *Intelligence Artificielle, mythes et réalités*, Colloque Grands Projets et Systemes Complexes, Arcachon, 18 Jun. 2018; *Intelligence Artificielle, mythes et réalités*, Journée lab. CRISTAL (U. Lille), Gand, 6 Jul. 2018; *Shallow and Deep learning at TAU*, Keynote at the DATAIA-JST International Symposium on Data Science and AI, 10 Jul. 2018; *Une brève introduction à l'Intelligence Artificielle et au rapport Villani*, Académie des Technologies, 10 Oct. 2018; *Une brève introduction à l'Intelligence Artificielle*, Open-Lab PSA, 11 Oct. 2018; *A brief introduction to AI and Deep Learning*, Toulouse Symposium on Deep Learning, 18 oct. 2018; *A brief introduction to AI and Deep Learning*, PSA Stellab Seminar, Paris, 14 nov. 2018; *IA, une stratégie française : le rapport Villani*, Université Franco Italienne, Università Italo Francese, annual seminar, Paris, 21 Nov. 2018; *Some issues with Deep Learning*, Inria / Nokia-Bell Labs seminar, 27 Nov. 2018; *L'Intelligence Artificielle hier, aujourd'hui et demain*, 29e journée CASCIMODOT, Orléans, 12 Dec. 2018.

Paola Tubaro *Un champ de mines? Éthique, droit et politique dans la recherche sur les réseaux sociaux*, conférence MARAMI (Modèles & Analyse des Réseaux : Approches Mathématiques & Informatiques), Avignon, 17 Oct. 2018; *Modèles multi-agents et simulation*, Ecole thématique CNRS "Réseaux et complexité", 28 Sept. 2018; *Microworking in France: an inquiry into the human labour that makes AI possible*, OECD, Paris, 5 Dec. 2018; *Micro-work, artificial intelligence and the automotive industry*, Sant'Anna School of Advnced Studies, Pisa, 31 May 2018; *Online platform labor*, EIT Digital, Rennes, 14 Nov. 2018; *Les promesses et les périls du travail sur plateformes*, France Stratégie, 6 July 2018; *Faut-il interdire la parole problématique en ligne ?*, GEPS medical conference, Montpellier, 12 Jan. 2018

10.1.5. Leadership within the Scientific Community

Isabelle Guyon President and co-founder of ChaLearn, a non-for-profit organization dedicated to the organization of challenge.

Marc Schoenauer Chair of ACM-SIGEVO (Special Interest Group on Evolutionary Computation), 2015-2017, re-elected July 2017 (2-years term); Founding President (since 2015) of SPECIES (Society for the Promotion of Evolutionary Computation In Europe and Surroundings), that organizes the yearly series of conferences *EvoStar*.

Michèle Sebag Elected Chair of Steering Committee, ECML-PKDD; Head of the Research Programme, Institut de Convergence DataIA.

Paola Tubaro Convenor of the Social Network Analysis Group of British Sociological Association; co-founder of European Network on Digital Labor.

10.1.6. Scientific Expertise

Cécile Germain Evaluator for the H2020-2016-CNECT program; member of the DFG review panel within Germany's excellence strategy selection process.

Marc Schoenauer Member, Villani Mission on Artificial Intelligence [50] (see also the AIforHumanity Web Site); Conseil scientifique, MoveInSalcay platform, coordinated by Nokia-Bell Labs; Comité Scientifique IA, SCube (Scientipôle Savoirs & Société), Orsay; Scientific Committee, TrackML (see Section 7.6); Comité de sélection, Chaire ABEONA-ENS "Biais et Equité en IA"; Conseil Scientifique, Fondation de Recherche pour l'Aéronautique et l'Espace (FRAE).

Michèle Sebag Jury de sélection, LRI; ENS-Lyon; LIX-Ecole Polytechnique; LORIA; Univ. Dortmund; Univ. Liège. Evaluation NSERC, Canada.

10.1.7. Research Administration

Cécile Germain University officer for scientific computing; member of the Board of the Lidex *Center for Data Science*; member of the scientific council of faculty of Medicine (UPsud).

Isabelle Guyon Representative of UPSud in the DataIA *Institut de Convergence* Program Committee, University of Paris-Saclay.

Marc Schoenauer Deputy Scientific Director of Inria (in French, Directeur Scientifique Adjoint, DSA), in charge of AI.

Michele Sebag Deputy director of LRI, CNRS UMR 8623; elected member of the Research Council of Univ. Paris-Saclay; member of the STIC department council of Univ. Paris-Saclay; member of the Scientific Council of Labex AMIES, Applications des Mathématiques ds l'Industrie, l'Entreprise et la Société; member of the Scientific Council of IRT System'X; member of the CSFRS (Conseil supérieur de la formation et de la recherche stratégique).

Paola Tubaro Representative of CNRS in the DataIA *Institut de Convergence* Program Committee, University of Paris-Saclay; member of the Board, Maison des Sciences de l'Homme Paris-Saclay.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence : Philippe Caillou, Computer Science for students in Accounting and Management, 192h, L1, IUT Sceaux, Univ. Paris Sud.

Licence : Aurélien Decelle, Computer Architecture, 26h, L2, Univ. Paris-Sud.

Licence : Aurélien Decelle, Introdution to Machine Learning, 57h, L2, Univ. Paris-Sud.

Licence : Aurélien Decelle, Object-oriented programming , 26h, L2, Univ. Paris-Sud.

Licence : Aurélien Decelle, Computer Architecture, 26h, L3, Univ. Paris-Sud.

Licence : François Landes, Mathematics for Computer Scientists, 51h, L2, Univ. Paris-Sud.

Licence : François Landes, Machine Learning and Artificial Life, 48h, L2, Univ. Paris-Sud.

Licence and Polytech : Cécile Germain, Computer Architecture

Licence : Isabelle Guyon, Project: Resolution of mini-challenges (created by M2 students), L2, Univ. Paris-Sud.

Master : François Landes, Machine Learning, 22h, M1 Polytech, U. Paris-sud.

Master : Guillaume Charpiat and Victor Berger, Advanced Machine Learning, 34h, M2 Recherche, Centrale-Supélec.

Master : Guillaume Charpiat, Introduction to Deep Learning, 6h, M2 Recherche, Telecom.

Master : Aurélien Decelle, Machine Learning, 26h, M1, Univ. Paris-Sud.

Master : Aurélien Decelle, Probability and statistics, 26h, M1, Univ. Paris-Sud.

Master : Cécile Germain, Parallel Programming

Master : Isabelle Guyon, Project: Creation of mini-challenges, M2, Univ. Paris-Sud.

Master : Michèle Sebag, Machine Learning, 12h; Deep Learning, 9h; Reinforcement Learning, 12h; M2 Recherche, U. Paris-sud. Summer School Deep Learning, Genova (Italy), 5h.

Master : François Landes, Machine Learning, 9h, M2 Recherche, U. Paris-sud.

Master : Paola Tubaro, Start -up project for engineering students, 24h, Telecom ParisTech.

Master : Paola Tubaro, Sociology of social networks, 24h, M2, EHESS/ENS.

Master : Flora Jay, Population Genetics, 10h, M2, Univ. Paris-Sud.

Doctorate: Paola Tubaro, Research Methods, 12h, University of Insubria, Italy.

10.2.2. Supervision

PhD François GONARD, *Cold-start recommendation : from Algorithm Portfolios to Job Applicant Matching*, Université Paris-Saclay, May 2018

PhD Thomas SCHMITT, *Collaborative Matching of Job Openings and Job Seekers*, Université Paris-Saclay, June 2018

PhD Hoang M. LUONG, *Squaring the Circle in Modelling Corporate Governance, Market Structure and Innovation: A Tobin's Q Approach to R&D Investment when Network Effects Are Present*, Paola Tubaro, with M. Ugur and S. Gorgoni, University of Greenwich, London, UK, Oct. 2018.

PhD in progress Eléonore BARTENLIAN, *Deep Learning pour le traitement du signal*, 1/10/2018, Michèle Sebag and Frédéric Pascal (Centrale-Supélec)

PhD in progress Victor BERGER, *Variational Anytime Simulator*, 1/10/2017, Michèle Sebag

PhD in progress Leonard BLIER, *Vers une architecture stable pour les systèmes d'apprentissage par renforcement*, 1/09/2018, Yann Ollivier (Facebook AI Research, Paris) and Marc Schoenauer

PhD in progress Tony BONNAIRE, *Reconstruction de la toile cosmique*, from 1/10/2018, Nabila Aghanim (Institut d'Astrophysique Spatiale) and Aurélien Decelle

PhD in progress Benjamin DONNOT, *Optimisation et méthodes d'apprentissage pour une conduite robuste et efficace du réseau électrique par anticipation sur base de parades topologiques.*, 1/09/2015, Isabelle Guyon and Antoine Marot (RTE)

PhD in progress Balthazar DONON, *Apprentissage par renforcement pour une conduite stratégique du système électrique*, 1/10/2018, Isabelle Guyon and Antoine Marot (RTE)

PhD in progress Guillaume DOQUET, *ML Algorithm Selection and Domain Adaptation*, 1/09/2015, Michele Sebag

PhD in progress Victor ESTRADE *Robust domain-adversarial learning, with applications to High Energy Physics,* 01/10/2016, Cécile Germain and Isabelle Guyon.

PhD in progress Loris FELARDOS, *Neural networks for molecular dynamics simulations*, 1/10/2018, Guillaume Charpiat, Jérôme Hénin (IBPC) and Bruno Raffin (InriAlpes)

PhD in progress Giancarlo FISSORE, *Statistical physics analysis of generative models*, 1/10/2017, Aurélien Decelle and Cyril Furtlehner

PhD in progress Julien GIRARD, *Vérification et validation des techniques d'apprentissage automatique*, 1/10/2018, Zakarian Chihani (CEA) and Guillaume Charpiat

PhD in progress Diviyan KALAINATHAN, *Causal models and quality of life at work*, 1/10/2017, Michèle Sebag and Isabelle Guyon

PhD in progress Zhengying LIU, *Automation du design des reseaux de neurones profonds*, 1/10/2017, Isabelle Guyon

PhD in progress Nizam MAKDOUD, *Motivations intrinsèques en apprentissage par renforcement. Application à la recherche de failles de sécurité*, 1/02/2018, Marc Schoenauer and Jérôme Kodjabachian (Thalès ThereSIS, Palaiseau).

PhD in progress Marc NABHAN, *Sûreté de fonctionnement d'un véhicule autonome - évaluation des fausses détections au travers d'un profil de mission réduit*, 1/10/2017, Marc Schoenauer and Hiba Hage (Renault)

PhD in progress Anna PIAZZA, *Inter-Organisational Relationships and Organisational Performance: Network Analysis Applications to a Health Care System*, 01/09/2014, Paola Tubaro, with F. Pallotti and A. Lomi, at the University of Greenwich, London, UK

PhD in progress Adrian POL *Machine Learning Anomaly Detection, with application to CMS Data Quality Monitoring,* 01/10/2016, Cécile Germain.

PhD in progress Herilalaina RAKOTOARISON, *Automatic Algorithm Configuration for Power Grid Optimization*, 1/10/2017, Marc Schoenauer and Michèle Sebag

PhD in progress Théophile SANCHEZ, *Reconstructing the past: deep learning for population genetics*, 1/10/2017, Guillaume Charpiat and Flora Jay

PhD in progress Lisheng SUN, *Apprentissage Automatique: Vers une analyse de données automatisé*, 1/10/2016, Isabelle Guyon and Michèle Sebag

PhD in progress Corentin TALLEC, *Reinforcement Learning and Recurrent Neural Networks: Dynamical approaches*, 1/10/2016, Yann Ollivier

PhD in progress Marion ULLMO, *Detection et classification de la toile et des filaments cosmiques*, from 1/10/2018, Nabila Aghanim (Institut d'Astrophysique Spatiale) and Aurélien Decelle

PhD in progress Pierre WOLINSKI, *Learning the Architecture of Neural Networks*, 1/9/2016, Yann Ollivier (Facebook AI Research, Paris) and Guillaume Charpiat

10.2.3. Juries

- Guillaume Charpiat As a member of the "Commission Scientifique" of Inria Saclay: selection committee for post-docs and PhD students hiring; jury of the Gilles Khan PhD prize (SIF); jury for a "Maître de conférence" position at Université Paris-Sud.
- Cécile Germain PhD jury for Mehdi Cherti 26/01/18, Wenjie ZHENG, 13/06/2018; jury of the Telecom PhD prize 23/03/2018
- Isabelle Guyon PhD jury for Giorgos Borboudakis, 21 Nov. 2018, University of Crete, Heraklion, Greece; Master thesis jury for Marvin Lerousseau, 28 June 2018, INP Grenoble; Master thesis jury for Adrien Pavao, 5 Sep. 2018, U.P-Sud, Orsay.
- Michèle Sebag PhD reviewer: Antonio Vergari (Univ. Bari, Italy); Gaétant Hadjeres, LIX, Ecole Polytechnique; Romain Warlop, Univ. Lille. Pdt Jury: Hafiz Tiomoko, Centrale-Supélec; Stanislas Chambon, Telecom.
- Marc Schoenauer PhD jury for Wen SUN, Université d'Angers, 29/11/2018;
- Paola Tubaro PhD jury for Victorien Barbet, Aix-Marseille Université, 13/12/2018.

10.3. Popularization

10.3.1. Articles and contents

Isabelle Guyon

- Pionnière : Isabelle Guyon, professeur à l'université de Paris-Saclay, 7 Feb. 2018, Usine Nouvelle
- Isabelle Guyon veut démocratiser l'intelligence artificielle, Apr. 2018, Le Monde
- Portrait : Isabelle Guyon, sélectionneuse d'algorithmes, 16 May 2018, Industries et Technologie

Paola Tubaro *Notre vie privée, un concept négociable* [52], Antonio Casilli and Paola Tubaro, in *Le Monde*, Jan. 2018.

10.3.2. Interventions

Guillaume Charpiat Séminaire d'Actualité Critique, *L'intelligence artificielle, entre fantasmes et recherche*, ENS Paris, 5 Apr. 2018.

Isabelle Guyon *Codalab, createur de defis*, UniThe ou Café, Inria Saclay, 5 Apr. 2018

Marc Schoenauer

- Congrès Européen Ethique et Gouvernance 2018, OCDE, Paris, 30 Mar. 2018;

- Audition publique sur l'Intelligence Artificielle et le travail, Comité économique et social européen (CESE), Bruxelles, 30 May 2018;
- Public debate with Jean-Marc David, Convention Systematic, 5 Jun. 2018; Conférence Legrain, ENS Paris, 5 Jul. 2018;
- Prix Design et Science de l'Université Paris-Saclay, 11 Oct. 2018.

Michèle Sebag Usine Nouvelle; Podcast Science et Avenir, discussion avec Cédric Villani.

11. Bibliography

Major publications by the team in recent years

[1] C. ADAM-BOURDARIOS, G. COWAN, C. GERMAIN-RENAUD, I. GUYON, B. KÉGL, D. ROUSSEAU.*The Higgs Machine Learning Challenge*, in "Journal of Physics: Conference Series", December 2015, vol. 664, n° 7 [*DOI :* 10.1088/1742-6596/664/7/072015], https://hal.inria.fr/hal-01745998

[2] L. DA COSTA, Á. FIALHO, M. SCHOENAUER, M. SEBAG.*Adaptive Operator Selection with Dynamic Multi-Armed Bandits*, in "Proc. Genetic and Evolutionary Computation Conference (GECCO)", ACM, 2008, p. 913-920, ACM-SIGEVO 10-years Impact Award [*DOI :* 10.1145/1389095.1389272], https://hal.inria.fr/inria-00278542

[3] C. FURTLEHNER, A. DECELLE.*Cycle-based Cluster Variational Method for Direct and Inverse Inference*, in "Journal of Statistical Physics", August 2016, vol. 164, n° 3, p. 531–574, https://hal.inria.fr/hal-01214155

[4] S. GELLY, M. SCHOENAUER, M. SEBAG, O. TEYTAUD, L. KOCSIS, D. SILVER, C. SZEPESVARI.*The Grand Challenge of Computer Go: Monte Carlo Tree Search and Extensions*, in "Communications- ACM", 2012, vol. 55, n° 3, p. 106-113, https://hal.inria.fr/hal-00695370

[5] O. GOUDET, D. KALAINATHAN, P. CAILLOU, D. LOPEZ-PAZ, I. GUYON, M. SEBAG.*Learning Functional Causal Models with Generative Neural Networks*, in "Explainable and Interpretable Models in Computer Vision and Machine Learning", Springer Series on Challenges in Machine Learning, Springer International Publishing, 2018, https://arxiv.org/abs/1709.05321 [*DOI :* 10.1007/978-3-319-98131-4], https://hal.archives-ouvertes.fr/hal-01649153

[6] T. LUCAS, C. TALLEC, J. VERBEEK, Y. OLLIVIER.*Mixed batches and symmetric discriminators for GAN training*, in "ICML - 35th International Conference on Machine Learning", Stockholm, Sweden, July 2018, https://hal.inria.fr/hal-01791126

[7] E. MAGGIORI, Y. TARABALKA, G. CHARPIAT, P. ALLIEZ.*Convolutional Neural Networks for Large-Scale Remote Sensing Image Classification*, in "IEEE Transactions on Geoscience and Remote Sensing", 2017, vol. 55, n° 2, p. 645-657, https://hal.inria.fr/hal-01369906

[8] M. MISIR, M. SEBAG.*Alors: An algorithm recommender system*, in "Artificial Intelligence", 2017, vol. 244, p. 291-314, Published on-line Dec. 2016, https://hal.inria.fr/hal-01419874

[9] Y. OLLIVIER, L. ARNOLD, A. AUGER, N. HANSEN.*Information-Geometric Optimization Algorithms: A Unifying Picture via Invariance Principles*, in "Journal of Machine Learning Research", 2017, vol. 18, n° 18, p. 1-65, https://hal.inria.fr/hal-01515898

[10] X. ZHANG, C. FURTLEHNER, C. GERMAIN-RENAUD, M. SEBAG.*Data Stream Clustering with Affinity Propagation*, in "IEEE Transactions on Knowledge and Data Engineering", 2014, vol. 26, n^{o} 7, 1, https://hal.inria.fr/hal-00862941

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] F. GONARD.*Cold-start recommendation : from Algorithm Portfolios to Job Applicant Matching*, Université Paris-Saclay, May 2018, https://tel.archives-ouvertes.fr/tel-01825220

[12] T. SCHMITT.*Collaborative Matching of Job Openings and Job Seekers*, Université Paris-Saclay, June 2018, https://tel.archives-ouvertes.fr/tel-01886623

Articles in International Peer-Reviewed Journal

[13] G. BASTIN, P. TUBARO.*Le moment big data des sciences sociales*, in "Revue française de sociologie", September 2018, vol. 59, n^{o} 3, p. 375-394 [*DOI :* 10.3917/RFS.593.0375], https://hal.archives-ouvertes.fr/hal-01885416

[14] K. CAYE, F. JAY, O. MICHEL, O. FRANÇOIS.*Fast Inference of Individual Admixture Coefficients Using Geographic Data*, in "Annals Of Applied Statistics", 2018, https://hal.archives-ouvertes.fr/hal-01676712

[15] A. DECELLE, G. FISSORE, C. FURTLEHNER.*Thermodynamics of Restricted Boltzmann Machines and Related Learning Dynamics*, in "Journal of Statistical Physics", January 2018, vol. 172, n^{o} 6, p. 1576-1608 [*DOI :* 10.1007/S10955-018-2105-Y], https://hal.inria.fr/hal-01675310

[16] S. ESCALERA, X. BARÓ, I. GUYON, H. J. ESCALANTE.*Guest Editorial: Apparent Personality Analysis*, in "IEEE Transactions on Affective Computing", January 2018, vol. 9, n^{o} 3, p. 299-302 [*DOI :* 10.1007/978-3-319-98131-4], https://hal.inria.fr/hal-01991636

[17] S. ESCALERA, X. BARÓ, I. GUYON, H. J. ESCALANTE, G. TZIMIROPOULOS, M. VALSTAR, M. PANTIC, J. COHN, T. KANADE.*Guest Editorial: The Computational Face*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", November 2018, vol. 40, n^{o} 11, p. 2541-2545 [*DOI :* 10.1109/TPAMI.2018.2869610], https://hal.inria.fr/hal-01991654

[18] F. PALLOTTI, P. TUBARO, A. A. CASILLI, T. W. VALENTE.*"You see yourself like in a mirror": The effects of internet-mediated personal networks on body image and eating disorders*, in "Health Communication", 2018, vol. 33, n^{o} 9, p. 1166-1176, Published online on 6 July 2017 [*DOI :* 10.1080/10410236.2017.1339371], https://hal.archives-ouvertes.fr/hal-01520138

[19] N. SHAH, B. TABIBIAN, K. MUANDET, I. GUYON, U. VON LUXBURG.*Design and Analysis of the NIPS 2016 Review Process*, in "Journal of Machine Learning Research", 2018, vol. 19, p. 1 - 34, https://hal.inria.fr/hal-01991580

[20] P. TAILLANDIER, B. GAUDOU, A. GRIGNARD, Q.-N. HUYNH, N. MARILLEAU, P. CAILLOU, D. PHILIPPON, A. DROGOUL.*Building, composing and experimenting complex spatial models with the GAMA platform*, in "GeoInformatica", December 2018, https://hal.inria.fr/hal-01975984

International Conferences with Proceedings

[21] M.-L. Cauwet, J. Decock, J. Liu, O. Teytaud. *Direct Model Predictive Control: A Theoretical and Numerical Analysis*, in "PSCC 2018 - XX Power Systems Computation Conference", Dublin, Ireland, June 2018, https://hal.inria.fr/hal-01701623

[22] M.-L. Cauwet, O. Teytaud. *Surprising strategies obtained by stochastic optimization in partially observable games*, in "CEC 2018 - IEEE Congress on Evolutionary Computation", Rio de Janeiro, Brazil, July 2018, p. 1-8, https://arxiv.org/abs/1807.01877 , https://hal.inria.fr/hal-01829721

[23] B. Donnot, I. Guyon, M. Schoenauer, A. Marot, P. Panciatici. *Anticipating contingengies in power grids using fast neural net screening*, in "IEEE International Joint Conference on Neural Networks", Rio de Janeiro, Brazil, Proc. 2018 International Joint Conference on Neural Networks (IJCNN), IEEE, July 2018, https://arxiv.org/abs/1805.02608 , https://hal.archives-ouvertes.fr/hal-01783669

[24] B. Donnot, I. Guyon, M. Schoenauer, A. Marot, P. Panciatici. *Fast Power system security analysis with Guided Dropout*, in "26th European Symposium on Artificial Neural Networks", Bruges, Belgium, Electronic Proceedings ESANN 2018, April 2018, https://arxiv.org/abs/1801.09870 , https://hal.archives-ouvertes.fr/hal-01695793

[25] V. Estrade, C. Germain, I. Guyon, D. Rousseau. *Systematics aware learning: a case study in High Energy Physics*, in "ESANN 2018 - 26th European Symposium on Artificial Neural Networks", Bruges, Belgium, April 2018, https://hal.inria.fr/hal-01715155

[26] S. Giffard-Roisin, D. Gagne, A. Boucaud, B. Kégl, M. Yang, G. Charpiat, C. Monteleoni. *The 2018 Climate Informatics Hackathon: Hurricane Intensity Forecast*, in "8th International Workshop on Climate Informatics", Boulder, CO, United States, PROCEEDINGS OF THE 8TH INTERNATIONAL WORKSHOP ON CLIMATE INFORMATICS: CI 2018, September 2018, 4, https://hal.archives-ouvertes.fr/hal-01924336

[27] N. Girard, G. Charpiat, Y. Tarabalka. *Aligning and Updating Cadaster Maps with Aerial Images by Multi-Task, Multi-Resolution Deep Learning*, in "Asian Conference on Computer Vision (ACCV)", Perth, Australia, Proceedings, December 2018, https://hal.archives-ouvertes.fr/hal-01923568

[28] T. Lucas, C. Tallec, J. Verbeek, Y. Ollivier. *Mixed batches and symmetric discriminators for GAN training*, in "ICML - 35th International Conference on Machine Learning", Stockholm, Sweden, July 2018, https://hal.inria.fr/hal-01791126

[29] G. Ramírez, E. Villatoro, B. Ionescu, H. J. Escalante, S. Escalera, M. Larson, H. Müller, I. Guyon. *Overview of the Multimedia Information Processing for Personality and Social Networks Analysis Contest*, in "International Conference on Pattern Recognition (ICPR)", Beijing, China, T. Tan, J. Kittler, A. Jain (editors), IEEE, August 2018, p. 127-139, https://hal.inria.fr/hal-01991599

[30] D. Rousseau, S. Amrouche, P. Calafiura, V. Estrade, S. Farrell, C. Germain, V. Gligorov, T. Golling, H. Gray, I. Guyon, M. Hushchyn, V. Innocente, M. Kiehn, A. Salzburger, A. Ustyuzhanin, J.-R. V. Vlimant, Y. Yilmaz. *The TrackML challenge*, in "NIPS 2018 - 32nd Annual Conference on Neural Information Processing Systems", Montreal, Canada, December 2018, p. 1-23, https://hal.inria.fr/hal-01745714

[31] A. SCHOENAUER SEBAG, L. HEINRICH, M. SCHOENAUER, M. SEBAG, L. WU, S. ALTSCHULER.*Multi-Domain Adversarial Learning*, in "ICLR'19 - Seventh annual International Conference on Learning Representations", New Orleans, United States, Proc. ICLR 2019, Tara Sainath, May 2019, https://hal.inria.fr/hal-01968180

[32] C. TALLEC, Y. OLLIVIER.*Can recurrent neural networks warp time?*, in "International Conference on Learning Representation 2018", Vancouver, France, April 2018, https://hal.inria.fr/hal-01812064

[33] P. TUBARO.*Emergent relational structures at a "sharing economy" festival*, in "The 7th International Conference on Complex Networks and Their Applications", Cambridge, United Kingdom, Springer, December 2018, vol. 2, p. 559-571, https://hal.archives-ouvertes.fr/hal-01947911

[34] A. ZAMPIERI, G. CHARPIAT, N. GIRARD, Y. TARABALKA.*Multimodal image alignment through a multi-scale chain of neural networks with application to remote sensing*, in "European Conference on Computer Vision (ECCV)", Munich, Germany, V. FERRARI, M. HEBERT, C. SMINCHISESCU, Y. WEISS (editors), Computer Vision – ECCV 2018, Springer International Publishing, September 2018, p. 679-696, https://hal.inria.fr/hal-01849389

Conferences without Proceedings

[35] S. GIFFARD-ROISIN, M. YANG, G. CHARPIAT, B. KÉGL, C. MONTELEONI.*Deep Learning for Hurricane Track Forecasting from Aligned Spatio-temporal Climate Datasets*, in "Modeling and decision-making in the spatiotemporal domain NIPS workhop", Montréal, Canada, December 2018, https://hal.archives-ouvertes.fr/hal-01905408

[36] S. GIFFARD-ROISIN, M. YANG, G. CHARPIAT, B. KÉGL, C. MONTELEONI.*Fused Deep Learning for Hurricane Track Forecast from Reanalysis Data*, in "Climate Informatics Workshop Proceedings 2018", Boulder, United States, September 2018, https://hal.archives-ouvertes.fr/hal-01851001

[37] Z. LIU, O. BOUSQUET, A. ELISSEEFF, S. ESCALERA, I. GUYON, J. JACQUES, A. PAVAO, D. SILVER, L. SUN-HOSOYA, S. TREGUER, W.-W. TU, J. WANG, Q. YAO.*AutoDL Challenge Design and Beta Tests-Towards automatic deep learning*, in "CiML workshop @ NIPS2018", Montreal, Canada, December 2018, https://hal.archives-ouvertes.fr/hal-01906226

[38] J. G. MADRID, H. JAIR ESCALANTE, E. F. MORALES, W.-W. TU, Y. YU, L. SUN-HOSOYA, I. GUYON, M. SEBAG.*Towards AutoML in the presence of Drift: first results*, in "Workshop AutoML 2018 @ ICML/IJCAI-ECAI", Stockholm, Sweden, Pavel Brazdil, Christophe Giraud-Carrier, and Isabelle Guyon, July 2018, https://hal.inria.fr/hal-01966962

[39] A. MAROT, B. DONON, I. GUYON, B. DONNOT.*Learning To Run a Power Network Competition*, in "CiML Workshop, NeurIPS", Montréal, Canada, December 2018, https://hal.archives-ouvertes.fr/hal-01968295

[40] H. RAKOTOARISON, M. SEBAG.*AutoML with Monte Carlo Tree Search*, in "Workshop AutoML 2018 @ ICML/IJCAI-ECAI", Stockholm, Sweden, Pavel Brazdil, Christophe Giraud-Carrier, and Isabelle Guyon, July 2018, https://hal.inria.fr/hal-01966957

[41] H. RICHARD, A. PINHO, B. THIRION, G. CHARPIAT.*Optimizing deep video representation to match brain activity*, in "CCN 2018 - Conference on Cognitive Computational Neuroscience", Philadelphia, United States, September 2018, https://arxiv.org/abs/1809.02440 , https://hal.archives-ouvertes.fr/hal-01868735

[42] L. SUN-HOSOYA, I. GUYON, M. SEBAG.*Lessons learned from the AutoML challenge*, in "Conférence sur l'Apprentissage Automatique 2018", Rouen, France, June 2018, https://hal.inria.fr/hal-01811454

[43] C. TALLEC, Y. OLLIVIER.*Unbiased online recurrent optimization*, in "International Conference On Learning Representation", Vancouver, Canada, April 2018, https://hal.archives-ouvertes.fr/hal-01972587

Scientific Books (or Scientific Book chapters)

[44] O. GOUDET, D. KALAINATHAN, P. CAILLOU, D. LOPEZ-PAZ, I. GUYON, M. SEBAG.*Learning Functional Causal Models with Generative Neural Networks*, in "Explainable and Interpretable Models in Computer Vision and Machine Learning", Springer Series on Challenges in Machine Learning, Springer International Publishing, 2018, https://arxiv.org/abs/1709.05321 [*DOI :* 10.1007/978-3-319-98131-4], https://hal.archives-ouvertes.fr/hal-01649153

[45] I. GUYON, L. SUN-HOSOYA, M. BOULLÉ, H. J. ESCALANTE, S. ESCALERA, Z. LIU, D. JAJETIC, B. RAY, M. SAEED, M. SEBAG, A. STATNIKOV, W.-W. TU, E. VIEGAS.*Analysis of the AutoML Challenge series 2015-2018*, in "AutoML: Methods, Systems, Challenges", F. HUTTER, L. KOTTHOFF, J. VANSCHOREN (editors), The Springer Series on Challenges in Machine Learning, Springer Verlag, 2018, https://hal.archives-ouvertes.fr/hal-01906197

Books or Proceedings Editing

[46] G. BASTIN, P. TUBARO (editors). *Big data, sociétés et sciences sociales*, Presses de Sciences Po, September 2018, vol. 59, n^{o} 3, https://hal.archives-ouvertes.fr/hal-01885457

[47] H. J. ESCALANTE, S. ESCALERA, I. GUYON, X. BARÓ, Y. GÜÇLÜTÜRK, U. GÜÇLÜ, M. A. J. VAN GERVEN (editors). *Explainable and interpretable models in computer vision and machine learning*, The Springer Series on Challenges in Machine Learning, Springer Verlag, January 2018 [*DOI :* 10.1007/978-3-319-98131-4], https://hal.inria.fr/hal-01991623

[48] E. LUTTON, P. LEGRAND, P. PARREND, N. MONMARCHÉ, M. SCHOENAUER (editors). *Artificial Evolution, 13th International Conference, Evolution Artificielle, EA 2017, Paris, France, October 25-27, 2017,Revised Selected Papers*, Artificial Evolution, Springer, March 2018, vol. LNCS - Lecture Notes in Computer Science, n^{o} 10764 [*DOI :* 10.1007/978-3-319-78133-4], https://hal.inria.fr/hal-01910443

Research Reports

[49] C. FURTLEHNER, J.-M. LASGOUTTES, A. ATTANASI, L. MESCHINI, M. PEZZULLA.*Spatio-temporal Probabilistic Short-term Forecasting on Urban Networks*, Inria Saclay, équipe TAU ; Inria de Paris, équipe RITS ; PTV-SISTeMA, December 2018, n^{o} RR-9236, https://hal.inria.fr/hal-01964270

Scientific Popularization

[50] C. VILLANI, M. SCHOENAUER, Y. BONNET, C. BERTHET, A.-C. CORNUT, F. LEVIN, B. RONDEPIERRE.*Donner un sens à l'intelligence artificielle : Pour une stratégie nationale et européenne*, Mission Villani sur l'intelligence artificielle, March 2018, https://hal.inria.fr/hal-01967551

Other Publications

[51] M. A. BARRY, V. BERTHIER, M.-C. CAMBOURIEUX, R. POLLÈS, B. D. WILTS, O. TEYTAUD, E. CENTENO, N. BIAIS, A. MOREAU.*Evolutionary algorithms converge towards evolved biological photonic structures*, August 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01856961

[52] A. A. CASILLI, P. TUBARO.*Notre vie privée, un concept négociable*, Le Monde, January 2018, https://hal.archives-ouvertes.fr/hal-01885452

[53] B. DONNOT.*Learning To Run a Power Network Competition - Poster*, December 2018, CiML Workshop, NeurIPS, Poster, https://hal.archives-ouvertes.fr/hal-01978748

[54] B. DONNOT, I. GUYON, Z. LIU, M. SCHOENAUER, A. MAROT, P. PANCIATICI.*Latent Surgical Interventions in Residual Neural Networks*, October 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01906170

[55] B. DONNOT, I. GUYON, A. MAROT, M. SCHOENAUER, P. PANCIATICI.*Optimization of computational budget for power system risk assessment*, May 2018, https://arxiv.org/abs/1805.01174 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01783685

[56] H. J. ESCALANTE, H. KAYA, A. A. SALAH, S. ESCALERA, Y. GUCLUTURK, U. GÜÇLÜ, X. BARÓ, I. GUYON, J. J. JUNIOR, M. MADADI, S. AYACHE, E. VIEGAS, F. GURPINAR, A. S. WICAKSANA, C. C. S. LIEM, M. A. J. VAN GERVEN, R. VAN LIER.*Explaining First Impressions: Modeling, Recognizing, and Explaining Apparent Personality from Videos*, January 2019, https://arxiv.org/abs/1802.00745 - Preprint submitted to IJCV, https://hal.inria.fr/hal-01991652

[57] D. KALAINATHAN, O. GOUDET, I. GUYON, D. LOPEZ-PAZ, M. SEBAG.*SAM: Structural Agnostic Model, Causal Discovery and Penalized Adversarial Learning*, August 2018, https://arxiv.org/abs/1803.04929 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01864239

[58] J. LEHMAN, J. CLUNE, D. MISEVIC, C. ADAMI, J. BEAULIEU, P. J. BENTLEY, S. BERNARD, G. BESLON, D. M. BRYSON, N. CHENEY, A. CULLY, S. DONCIEUX, F. C. DYER, K. O. ELLEFSEN, R. FELDT, S. FISCHER, S. FORREST, A. FRENOY, C. GAGNEÉ, L. LE GOFF, L. M. GRABOWSKI, B. HODJAT, L. KELLER, C. KNIBBE, P. KRCAH, R. E. LENSKI, H. LIPSON, R. MACCURDY, C. MAESTRE, R. MIIKKULAINEN, S. MITRI, D. E. MORIARTY, J.-B. MOURET, A. D. NGUYEN, C. OFRIA, M. PARIZEAU, D. PARSONS, R. T. PENNOCK, W. F. PUNCH, T. S. RAY, M. SCHOENAUER, E. SHULTE, K. SIMS, K. O. STANLEY, F. TADDEI, D. TARAPORE, S. THIBAULT, W. WEIMER, R. WATSON, J. YOSINKSI.*The Surprising Creativity of Digital Evolution: A Collection of Anecdotes from the Evolutionary Computation and Artificial Life Research Communities*, March 2018, https://arxiv.org/abs/1803.03453 - working paper or preprint, https://hal.inria.fr/hal-01735473

[59] S. LIÉGEOIS, O. FRANÇOIS, F. JAY.*Dimension reduction adapted to paleogenomics*, September 2018, Paris-Saclay Junior Conference on Data Science and Engineering, Poster, https://hal.archives-ouvertes.fr/hal-01978650

[60] A. A. POL, G. CERMINARA, C. GERMAIN, M. PIERINI, A. SETH.*Detector monitoring with artificial neural networks at the CMS experiment at the CERN Large Hadron Collider*, January 2019, https://arxiv.org/abs/1808.00911 - working paper or preprint, https://hal.inria.fr/hal-01976256

[61] D. ROUSSEAU, S. AMROUCHE, P. CALAFIURA, V. ESTRADE, S. FARRELL, C. GERMAIN, V. GLIGOROV, T. GOLLING, H. GRAY, I. GUYON, M. HUSHCHYN, V. INNOCENTE, M. KIEHN, A. SALZBURGER, A.

USTYUZHANIN, J.-R. V. VLIMANT, Y. YILMAZ.*The TrackML Particle Tracking Challenge*, July 2018, The document describes the challenge data, task and organization, https://hal.inria.fr/hal-01680537

[62] L. SUN-HOSOYA, I. GUYON, M. SEBAG.*ActivMetaL: Algorithm Recommendation with Active Meta Learning*, September 2018, IAL 2018 workshop, ECML PKDD, Poster, https://hal.archives-ouvertes.fr/hal-01931262

[63] A. ZAMPIERI, G. CHARPIAT, Y. TARABALKA.*Coarse to fine non-rigid registration: a chain of scale-specific neural networks for multimodal image alignment with application to remote sensing*, February 2018, https://arxiv.org/abs/1802.09816 - working paper or preprint, https://hal.inria.fr/hal-01718263

References in notes

[64] I. GUYON ET AL. (editor). *Cause Effect Pairs in Machine Learning*, Springer, 2019, In preparation

[65] C. ADAM-BOURDARIOS, G. COWAN, C. GERMAIN, I. GUYON, B. KÉGL, D. ROUSSEAU.*How Machine Learning won the Higgs Boson Challenge*, in "Proc. European Symposium on ANN, CI and ML", 2016

[66] C. ADAM-BOURDARIOS, G. COWAN, C. GERMAIN-RENAUD, I. GUYON, B. KÉGL, D. ROUSSEAU.*The Higgs Machine Learning Challenge*, in "Journal of Physics: Conference Series", 2015, vol. 664, n° 7 [*DOI :* 10.1088/1742-6596/664/7/072015], https://hal.inria.fr/hal-01745998

[67] A. ALAHI, K. GOEL, V. RAMANATHAN, A. ROBICQUET, L. FEI-FEI, S. SAVARESE.*Social LSTM: Human Trajectory Prediction in Crowded Spaces*, in "2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)", 2016, p. 961–971

[68] D. J. AMIT, H. GUTFREUND, H. SOMPOLINSKY.*Statistical Mechanics of Neural Networks Near Saturation*, in "Annals of Physics", 1987, vol. 173, p. 30-67

[69] C. ANDERSON.*The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*, 2008, Wired Magazine, https://www.wired.com/2008/06/pb-theory/

[70] W. BANZHAF, P. NORDIN, R. KELLER, F. FRANCONE.*Genetic Programming — An Introduction On the Automatic Evolution of Computer Programs and Its Applications*, Morgan Kaufmann, 1998

[71] N. BELKHIR.*Per Instance Algorithm Configuration for Continuous Black Box Optimization*, Université Paris-Saclay, November 2017, https://hal.inria.fr/tel-01669527

[72] N. BELKHIR, J. DRÉO, P. SAVÉANT, M. SCHOENAUER.*Per instance algorithm configuration of CMA-ES with limited budget*, in "Proc. ACM-GECCO", 2017, p. 681-688, https://hal.inria.fr/hal-01613753

[73] I. BELLO, B. ZOPH, V. VASUDEVAN, Q. V. LE.*Neural Optimizer Search with Reinforcement Learning*, in "Proceedings of the 34th International Conference on Machine Learning, ICML", 2017, p. 459–468, http://proceedings.mlr.press/v70/bello17a.html

[74] S. BEN-DAVID, J. BLITZER, K. CRAMMER, A. KULESZA, F. PEREIRA, J. W. VAUGHAN.*A theory of learning from different domains*, in "Machine Learning", 2018, vol. 79, n° 1, p. 151–175, http://link.springer.com/10.1007/s10994-009-5152-4

[75] J. BERGSTRA, R. BARDENET, Y. BENGIO, B. KÉGL.*Algorithms for Hyper-Parameter Optimization*, in "25th Annual Conference on Neural Information Processing Systems (NIPS 2011)", J. SHAWE-TAYLOR ET AL. (editor), Advances in Neural Information Processing Systems, Neural Information Processing Systems Foundation, 2011, vol. 24

[76] E. BOURDU, M.-M. PÉRÉTIÉ, M. RICHER.*La qualité de vie au travail : un levier de compétitivité*, Presses des Mines, 2016

[77] J. BRUNA, S. MALLAT.*Invariant Scattering Convolution Networks*, in "IEEE Trans. Pattern Anal. Mach. Intell.", 2013, vol. 35, n^{o} 8, p. 1872–1886

[78] Y. CHOI, M. CHOI, M. KIM, J. HA, S. KIM, J. CHOO.*StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation*, in "CoRR", 2017, vol. abs/1711.09020, http://arxiv.org/abs/1711.09020

[79] J.-J. CHRISTOPHE, J. DECOCK, O. TEYTAUD.*Direct model predictive control*, in "European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN)", 2014, https://hal.inria.fr/hal-00958192

[80] R. COUILLET, M. DEBBAH.*Random matrix methods for wireless communications*, Cambridge University Press, 2011

[81] DAVID SILVER ET AL..*Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm*, in "CoRR", 2017, vol. abs/1712.01815, http://arxiv.org/abs/1712.01815

[82] M. DEMARIA, M. MAINELLI, L. K. SHAY, J. A. KNAFF, J. KAPLAN.*Further improvements to the statistical hurricane intensity prediction scheme (SHIPS)*, in "Weather and Forecasting", 2005, vol. 20, n^{o} 4, p. 531–543

[83] A. DECELLE, G. FISSORE, C. FURTLEHNER.*Spectral dynamics of learning in restricted Boltzmann machines*, in "EPL (Europhysics Letters)", 2017, vol. 119, n^{o} 6, 60001

[84] L. DINH, J. SOHL-DICKSTEIN, S. BENGIO.*Density estimation using Real NVP*, in "Int. Conf. on Learning Representations (ICLR)", 2017

[85] T. DURIEZ, S. L. BRUNTON, B. R. NOACK.*Introduction*, Springer International Publishing, 2017, p. 1–10

[86] D. K. DUVENAUD, D. MACLAURIN, J. AGUILERA-IPARRAGUIRRE, R. GÓMEZ-BOMBARELLI, T. HIRZEL, A. ASPURU-GUZIK, R. P. ADAMS.*Convolutional Networks on Graphs for Learning Molecular Fingerprints*, in "NIPS", 2015, p. 2224–2232

[87] M. EICKENBERG, A. GRAMFORT, G. VAROQUAUX, B. THIRION.*Seeing it all: Convolutional network layers map the function of the human visual system*, in "NeuroImage", 2016 [*DOI :* 10.1016/J.NEUROIMAGE.2016.10.001], https://hal.inria.fr/hal-01389809

[88] T. V. ERVEN, P. GRUNWALD, S. DE ROOIJ.*Catching up faster by switching sooner: a predictive approach to adaptive estimation with an application to the AIC-BIC dilemma: Catching up Faster*, in "Journal of the Royal Statistical Society: Series B (Statistical Methodology)", 2012, vol. 74, n^{o} 3, p. 361–417, http://doi.wiley.com/10.1111/j.1467-9868.2011.01025.x

[89] H. J. ESCALANTE, I. GUYON, S. ESCALERA, J. JACQUES, M. MADADI, X. BARÓ, S. AYACHE, E. VIEGAS, Y. GÜÇLÜTÜRK, U. GÜÇLÜ, M. A. J. VAN GERVEN, R. VAN LIER.*Design of an Explainable Machine Learning Challenge for Video Interviews*, in "IJCNN 2017 - 30th International Joint Conference on Neural Networks", Anchorage, AK, United States, Neural Networks (IJCNN), 2017 International Joint Conference on, IEEE, 2017, p. 1-8, https://hal.inria.fr/hal-01668386

[90] S. ESCALERA, X. BARÓ, H. J. ESCALANTE, I. GUYON.*ChaLearn looking at people: A review of events and resources*, in "2017 International Joint Conference on Neural Networks (IJCNN)", 2017, p. 1594-1601

[91] S. FALKNER, A. KLEIN, F. HUTTER.*BOHB: Robust and Efficient Hyperparameter Optimization at Scale*, in "Proceedings of the 35th International Conference on Machine Learning (ICML 2018)", 2018, p. 1436–1445

[92] M. FEURER, A. KLEIN, K. EGGENSPERGER, J. SPRINGENBERG, M. BLUM, F. HUTTER.*Efficient and Robust Automated Machine Learning*, in "Advances in Neural Information Processing Systems 28", C. CORTES, N. D. LAWRENCE, D. D. LEE, M. SUGIYAMA, R. GARNETT (editors), Curran Associates, Inc., 2015, p. 2962–2970, http://papers.nips.cc/paper/5872-efficient-and-robust-automated-machine-learning.pdf

[93] L. FRANCESCHI, P. FRASCONI, S. SALZO, R. GRAZZI, M. PONTIL.*Bilevel Programming for Hyperparameter Optimization and Meta-Learning*, in "Proc. 35th International Conference on Machine Learning, ICML", J. G. DY, A. KRAUSE (editors), JMLR Workshop and Conference Proceedings, JMLR.org, 2018, vol. 80, p. 1563–1572, http://proceedings.mlr.press/v80/franceschi18a.html

[94] C. FURTLEHNER, A. DECELLE.*Cycle-Based Cluster Variational Method for Direct and Inverse Inference*, in "Journal of Statistical Physics", 2016, vol. 164, n^o 3, p. 531–574

[95] C. FURTLEHNER, M. SEBAG, X. ZHANG.*Scaling analysis of affinity propagation*, in "Physical Review E", 2010, vol. 81, n^o 6, 066102

[96] M. GABRIÉ, A. MANOEL, C. LUNEAU, J. BARBIER, N. MACRIS, F. KRZAKALA, L. ZDEBOROVÁ.*Entropy and mutual information in models of deep neural networks*, in "CoRR", 2018, vol. abs/1805.09785, http://arxiv.org/abs/1805.09785

[97] Y. GANIN, E. USTINOVA, H. AJAKAN, P. GERMAIN, H. LAROCHELLE, F. LAVIOLETTE, M. MARCHAND, V. LEMPITSKY.*Domain-Adversarial Training of Neural Networks*, in "Journal of Machine Learning Research", 2016, vol. 17, n^o 59, p. 1-35

[98] K. GARIMELLA, G. D. F. MORALES, A. GIONIS, M. MATHIOUDAKIS.*Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of Bipartisanship*, in "WWW", ACM, 2018, p. 913–922

[99] J. GILES.*Computational social science: Making the links*, in "Nature - News", 2012, vol. 488, n^o 7412, p. 448-450

[100] F. GONARD, M. SCHOENAUER, M. SEBAG.*ASAP.V2 and ASAP.V3: Sequential optimization of an Algorithm Selector and a Scheduler*, in "Open Algorithm Selection Challenge 2017 ", Brussels, Belgium, Proceedings of Machine Learning Research, September 2017, vol. 79, p. 8-11, https://hal.inria.fr/hal-01659700

[101] I. GOODFELLOW, J. POUGET-ABADIE, M. MIRZA, B. XU, D. WARDE-FARLEY, S. OZAIR, A. COURVILLE, Y. BENGIO.*Generative Adversarial Nets*, in "Advances in Neural Information Processing Systems 27", 2014, p. 2672–2680

[102] I. GOODFELLOW, J. POUGET-ABADIE, M. MIRZA, B. XU, D. WARDE-FARLEY, S. OZAIR, A. COURVILLE, Y. BENGIO.*Generative Adversarial Nets*, in "NIPS 27", Z. GHAHRAMANI, M. WELLING, C. CORTES, N. D. LAWRENCE, K. Q. WEINBERGER (editors), Curran Associates, Inc., 2014, p. 2672–2680

[103] P. GRUNWALD.*The Minimum Description Length Principle*, MIT Press, 2007

[104] I. GUYON, C. F. ALIFERIS, G. F. COOPER, A. ELISSEEFF, J.-P. PELLET, P. SPIRTES, A. R. STATNIKOV.*Design and Analysis of the Causation and Prediction Challenge*, in "WCCI Causation and Prediction Challenge", JMLR W-CP, 2008, p. 1–33

[105] I. GUYON, K. BENNETT, G. CAWLEY, H. J. ESCALANTE, S. ESCALERA, T. K. HO, N. MACIA, B. RAY, M. SAEED, A. STATNIKOV.*Design of the 2015 ChaLearn AutoML challenge*, in "Proc. IJCNN", IEEE, 2015, p. 1–8

[106] I. GUYON, K. BENNETT, G. CAWLEY, H. J. ESCALANTE, S. ESCALERA, T. K. HO, N. MACIA, B. RAY, M. SAEED, A. STATNIKOV.*Design of the 2015 ChaLearn AutoML challenge*, in "Proc. IJCNN", IEEE, 2015, p. 1–8

[107] I. GUYON, L. BOTTOU, B. SCHÖLKOPF, A. STATNIKOV, E. VIEGAS, J. M. ROBINS.*NIPS workshop on causality*, 2013, http://clopinet.com/isabelle/Projects/NIPS2013/

[108] U. GÜÇLÜ, M. A. J. VAN GERVEN.*Deep Neural Networks Reveal a Gradient in the Complexity of Neural Representations across the Ventral Stream*, in "Journal of Neuroscience", 2015, vol. 35, n^{o} 27, p. 10005–10014 [*DOI :* 10.1523/JNEUROSCI.5023-14.2015], http://www.jneurosci.org/content/35/27/10005

[109] D. HERNÁNDEZ-LOBATO, P. MORALES-MOMBIELA, D. LOPEZ-PAZ, A. SUÁREZ.*Non-linear Causal Inference using Gaussianity Measures*, in "Journal of Machine Learning Research", 2016, vol. 17, n^{o} 28, p. 1-39, http://jmlr.org/papers/v17/14-375.html

[110] H. H. HOOS.*Programming by Optimization*, in "Commun. ACM", 2012, vol. 55, n^{o} 2, p. 70–80

[111] D. HOTHERSALL.*History of Psychology*, McGraw-Hill, 2004

[112] G. W. IMBENS, D. B. RUBIN.*Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*, Cambridge University Press, 2015

[113] D. KALAINATHAN, O. GOUDET, P. CAILLOU, M. SEBAG, P. TUBARO, E. BOURDU, T. WEIL.*Portraits de travailleurs*, Presses des Mines, 2017

[114] D. KALAINATHAN, O. GOUDET, P. CAILLOU, M. SEBAG, P. TUBARO.*Conditions objectives de travail et ressenti des individus : le rôle du management*, 2017, 12, Synthèse n. 14 de ”La Fabrique de l'Industrie”, https://hal.inria.fr/hal-01742592

[115] W. KAY, J. CARREIRA, K. SIMONYAN, B. ZHANG, C. HILLIER, S. VIJAYANARASIMHAN, F. VIOLA, T. GREEN, T. BACK, P. NATSEV, M. SULEYMAN, A. ZISSERMAN.*The Kinetics Human Action Video Dataset*, in "arXiv e-prints", May 2017, arXiv:1705.06950

[116] D. KINGMA, M. WELLING.*Auto-encoding variational Bayes.*, in "Int. Conf. on Learning Representations (ICLR)", 2014

[117] L. KOTTHOFF.*Algorithm Selection for Combinatorial Search Problems: A Survey*, in "Data Mining and Constraint Programming: Foundations of a Cross-Disciplinary Approach", C. BESSIERE, L. DE RAEDT, L. KOTTHOFF, S. NIJSSEN, B. O'SULLIVAN, D. PEDRESCHI (editors), Lecture Notes in Computer Science, Springer, 2016, p. 149–190

[118] A. KRIZHEVSKY, I. SUTSKEVER, G. E. HINTON.*ImageNet Classification with Deep Convolutional Neural Networks*, in "Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1", NIPS'12, 2012, p. 1097–1105

[119] D. LAZER, A. PENTLAND, L. ADAMIC, S. ARAL, A. L. BARABASI, D. BREWER, N. CHRISTAKIS, N. CONTRACTOR, J. FOWLER, M. GUTMANN, T. JEBARA, G. KING, M. MACY, D. ROY, M. VAN ALSTYNE.*Life in the network: the coming age of computational social science*, in "Science", 2009, vol. 323, n^{o} 5915, p. 721–723

[120] Y. LECUN.*Deep Learning est mort. Vive Differentiable Programming!*, 2018, Facebook Post, https://www.facebook.com/yann.lecun/posts/10155003011462143

[121] J. LEHMAN, K. O. STANLEY.*Abandoning Objectives: Evolution Through the Search for Novelty Alone*, in "Evol. Comput.", 2011, vol. 19, n^{o} 2, p. 189–223, http://dx.doi.org/10.1162/EVCO_a_00025

[122] J. Z. LIANG, E. MEYERSON, R. MIIKKULAINEN.*Evolutionary architecture search for deep multitask networks*, in "Proc. ACM-GECCO", H. E. AGUIRRE, K. TAKADAMA (editors), ACM, 2018, p. 466–473, http://doi.acm.org/10.1145/3205455.3205489

[123] H. LIPSON, J. B. POLLACK.*Automatic design and manufacture of robotic lifeforms*, in "Nature Letters", 2000, vol. 406, p. 974–978

[124] D. LOPEZ-PAZ, K. MUANDET, B. SCHÖLKOPF, I. O. TOLSTIKHIN.*Towards a Learning Theory of Cause-Effect Inference*, in "Proceedings of the 32nd International Conference on Machine Learning, ICML", F. R. BACH, D. M. BLEI (editors), 2015, p. 1452–1461, http://jmlr.org/proceedings/papers/v37/lopez-paz15.html

[125] C. LOUIZOS, M. REISSER, T. BLANKEVOORT, E. GAVVES, M. WELLING.*Relaxed Quantization for Discretized Neural Networks*, in "CoRR", 2018, vol. abs/1810.01875, http://arxiv.org/abs/1810.01875

[126] C. LOUIZOS, M. REISSER, T. BLANKEVOORT, E. GAVVES, M. WELLING.*Relaxed Quantization for Discretized Neural Networks*, in "Int. Conf. on Learning Representations (ICLR)", 2019

[127] E. MAGGIORI, Y. TARABALKA, G. CHARPIAT, P. ALLIEZ.*Can Semantic Labeling Methods Generalize to Any City? The Inria Aerial Image Labeling Benchmark*, in "IEEE International Symposium on Geoscience and Remote Sensing (IGARSS)", Fort Worth, United States, July 2017, https://hal.inria.fr/hal-01468452

[128] L. MESCHEDER, A. GEIGER, S. NOWOZIN.*Which Training Methods for GANs do actually Converge?*, in "International Conference on Machine Learning", 2018, p. 3481–3490, http://proceedings.mlr.press/v80/mescheder18a.html

[129] D. MOLCHANOV, A. ASHUKHA, D. VETROV.*Variational Dropout Sparsifies Deep Neural Networks*, in "ArXiv e-prints", January 2017

[130] M. MISIR, M. SEBAG.*Alors: An algorithm recommender system*, in "Artificial Intelligence", 2017, vol. 244, p. 291-314, Published on-line Dec. 2016, https://hal.inria.fr/hal-01419874

[131] C. O'NEILL.*Weapons of Math Destruction*, Crown Books, 2016

[132] J. PEARL.*Causality: Models, Reasoning, and Inference (2nd edition)*, Cambridge University Press, 2009

[133] J. PEARL.*Theoretical Impediments to Machine Learning With Seven Sparks from the Causal Revolution*, in "Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, WSDM", 2018, 3, https://doi.org/10.1145/3159652.3176182

[134] G. PLEISS, M. RAGHAVAN, F. WU, J. M. KLEINBERG, K. Q. WEINBERGER.*On Fairness and Calibration*, in "NIPS", 2017, p. 5684–5693

[135] M. RAISSI.*Deep Hidden Physics Models: Deep Learning of Nonlinear Partial Differential Equations*, in "JMLR", 2018, vol. 19, p. 1–24

[136] S. RAVANBAKHSH, J. SCHNEIDER, B. PÓCZOS.*Equivariance Through Parameter-Sharing*, in "International Conference on Machine Learning", 2017, p. 2892–2901, http://proceedings.mlr.press/v70/ravanbakhsh17a.html

[137] E. REAL, S. MOORE, A. SELLE, S. SAXENA, Y. L. SUEMATSU, J. T. Q. LE, A. KURAKIN.*Large-Scale Evolution of Image Classifiers*, in "Proc. ICML 2017", 2017, https://icml.cc/Conferences/2017/Schedule?type=Poster

[138] J. R. RICE.*The Algorithm Selection Problem*, in "Advances in Computers", 1976, vol. 15, p. 65 – 118 [*DOI :* 10.1016/S0065-2458(08)60520-3], http://www.sciencedirect.com/science/article/pii/S0065245808605203

[139] J. RISSANEN.*Information and Complexity in Statistical Modeling*, Information Science and Statistics, Springer-Verlag, 2007

[140] P. ROLET, M. SEBAG, O. TEYTAUD.*Boosting Active Learning to Optimality: a Tractable Monte-Carlo, Billiard-based Algorithm*, in "ECML", Bled, Slovenia, 2009, p. 302-317, https://hal.inria.fr/inria-00433866

[141] T. SANCHEZ, G. CHARPIAT, F. JAY.*SPI-DNA: End-to-end Deep Learning Approach for Demographic History Inference*, in "Paris-Saclay Junior Conference on Data Science and Engineering", Orsay, France, September 2017, https://hal.archives-ouvertes.fr/hal-01679385

[142] M. SCHMIDT, H. LIPSON.*Distilling free-form natural laws from experimental data*, in "Science", 2009, vol. 324, n° 5923, p. 81–85

[143] T. SCHMITT, F. GONARD, P. CAILLOU, M. SEBAG.*Language Modelling for Collaborative Filtering: Application to Job Applicant Matching*, in "29th ICTAI", 2017, p. 1-8

[144] S. SIGTIA, E. BENETOS, S. DIXON.*An End-to-End Neural Network for Polyphonic Piano Music Transcription*, in "IEEE/ACM Trans. Audio, Speech & Language Processing", 2016, vol. 24, n^o 5, p. 927–939

[145] J. SIRIGNANO, K. SPILIOPOULOS.*DGM: A deep learning algorithm for solving partial differential equations*, in "Journal of Computational Physics", 2018, vol. 375, p. 1339 – 1364 [*DOI :* 10.1016/J.JCP.2018.08.029], http://www.sciencedirect.com/science/article/pii/S0021999118305527

[146] I. STENGERS.*Another Science Is Possible*, Open Humanities Press, 2013

[147] D. H. STERN, R. HERBRICH, T. GRAEPEL.*Matchbox: large scale online bayesian recommendations*, in "WWW", ACM Press, 2010, 111 [*DOI :* 10.1145/1526709.1526725], http://portal.acm.org/citation.cfm?doid=1526709.1526725

[148] C. SUNSTEIN, R. THALER.*Nudge: Improving Decisions about Health, Wealth, and Happiness*, Yale University Press, 2008

[149] C. SZEGEDY, W. ZAREMBA, I. SUTSKEVER, J. BRUNA, D. ERHAN, I. J. GOODFELLOW, R. FERGUS.*Intriguing properties of neural networks*, in "CoRR", 2013, vol. abs/1312.6199, http://arxiv.org/abs/1312.6199

[150] C. TALLEC, Y. OLLIVIER.*Can recurrent neural networks warp time?*, in "International Conference on Learning Representations – ICLR", 2018, https://openreview.net/forum?id=SJcKhk-Ab

[151] N. TISHBY, N. ZASLAVSKY.*Deep learning and the information bottleneck principle*, in "2015 IEEE Information Theory Workshop (ITW)", 2015, p. 1-5

[152] S. TRIANTAFILLOU, I. TSAMARDINOS.*Constraint-based Causal Discovery from Multiple Interventions over Overlapping Variable Sets*, in "Journal of Machine Learning Research", 2015, vol. 16, p. 2147-2205

[153] J. VANSCHOREN, J. N. VAN RIJN, B. BISCHL, L. TORGO.*OpenML: networked science in machine learning*, in "SIGKDD Explorations", 2013, vol. 15, n^o 2, p. 49–60, https://arxiv.org/abs/1407.7722

[154] VOLODYMYR MNIH ET AL..*Human-level control through deep reinforcement learning*, in "Nature", 2015, vol. 518, n^o 7540, p. 529–533, https://doi.org/10.1038/nature14236

[155] J. WAN, S. ESCALERA, G. ANBARJAFARI, H. J. ESCALANTE, X. BARÓ, I. GUYON, M. MADADI, J. ALLIK, C. LIN, Y. XIE.*Results and Analysis of ChaLearn LAP Multi-modal Isolated and Continuous Gesture Recognition, and Real versus Fake Expressed Emotions Challenges*, in "International Conference on Computer Vision - ICCV 2017", 2017, https://hal.inria.fr/hal-01677974

[156] M. WEIMER, A. KARATZOGLOU, Q. LE, A. SMOLA.*CoFiRank - Maximum Margin Matrix Factorization for Collaborative Ranking*, in "Advances in Neural Information Processing Systems, 21st Annual Conference on Neural Information Processing Systems 2007", Vancouver, Canada, December 2007, p. 222-230, http://books.nips.cc/papers/files/nips20/NIPS2007_0612.pdf, https://hal.archives-ouvertes.fr/hal-00482740

[157] M. WELLING.*Intelligence per Kilowatthour*, 2018, Keynote talk @ICML, https://icml.cc/Conferences/2018/Schedule?showEvent=1866

[158] F. M. J. WILLEMS, Y. M. SHTARKOV, T. J. TJALKENS.*The context-tree weighting method: basic properties*, in "IEEE Transactions on Information Theory", 1995, vol. 41, n^{O} 3, p. 653-664

[159] X. ZHANG, C. FURTLEHNER, C. GERMAIN-RENAUD, M. SEBAG.*Data Stream Clustering with Affinity Propagation*, in "IEEE Transactions on Knowledge and Data Engineering", 2014, vol. 26, n^{O} 7, 1, https://hal.inria.fr/hal-00862941

[160] M. DE BEZENAC, A. PAJOT, P. GALLINARI.*Deep Learning for Physical Processes: Incorporating Prior Scientific Knowledge*, in "arXiv preprint arXiv:1711.07970", 2017

[161] A. VAN DEN OORD, S. DIELEMAN, H. ZEN, K. SIMONYAN, O. VINYALS, A. GRAVES, N. KALCHBRENNER, A. SENIOR, K. KAVUKCUOGLU.*WaveNet: A Generative Model for Raw Audio*, in "CoRR", 2016, vol. abs/1609.03499, http://arxiv.org/abs/1609.03499

Project-Team TOCCATA

Certified Programs, Certified Tools, Certified Floating-Point Computations

IN COLLABORATION WITH: Laboratoire de recherche en informatique (LRI)

IN PARTNERSHIP WITH:
Université Paris-Sud (Paris 11)

RESEARCH CENTER
Saclay - Île-de-France

THEME
Proofs and Verification

Table of contents

Project-Team TOCCATA

Creation of the Team: 2012 September 01, updated into Project-Team: 2014 July 01

Keywords:

Computer Science and Digital Science:

A2.1.1. - Semantics of programming languages
A2.1.4. - Functional programming
A2.1.6. - Concurrent programming
A2.1.10. - Domain-specific languages
A2.1.11. - Proof languages
A2.4.2. - Model-checking
A2.4.3. - Proofs
A6.2.1. - Numerical analysis of PDE and ODE
A7.2. - Logic in Computer Science
A7.2.1. - Decision procedures
A7.2.2. - Automated Theorem Proving
A7.2.3. - Interactive Theorem Proving
A7.2.4. - Mechanized Formalization of Mathematics
A8.10. - Computer arithmetic

Other Research Topics and Application Domains:

B5.2.2. - Railway
B5.2.3. - Aviation
B5.2.4. - Aerospace
B6.1. - Software industry
B9.5.1. - Computer science
B9.5.2. - Mathematics

1. Team, Visitors, External Collaborators

Research Scientists

Claude Marché [Team leader, Inria, Senior Researcher, HDR]
Sylvie Boldo [Inria, Senior Researcher, HDR]
Jean-Christophe Filliâtre [CNRS, Senior Researcher, HDR]
Guillaume Melquiond [Inria, Researcher]

Faculty Members

Sylvain Conchon [Univ Paris-Sud, Professor, HDR]
Andrei Paskevich [Univ Paris-Sud, Associate Professor]
Thibault Hilaire [Associate Professor, Délégation de l'Univ Pierre et Marie Curie]
Benjamin Farinier [Univ Paris-Sud, ATER, from Oct 2018]

Technical Staff

Sylvain Dailler [Inria]

PhD Students

Martin Clochard [Inria, until Mar 2018]

Albin Coquereau [École Nationale Supérieure de Techniques Avancées]
David Declerck [Univ Paris-Sud, until Aug 2018]
Florian Faissole [Inria]
Diane Gallois-Wong [Univ Paris-Sud]
Quentin Garchery [Univ Paris-Sud, from Oct 2018]
Mário Pereira [Grant Portuguese government]
Raphaël Rieu-Helft [CIFRE TrustInSoft]
Mattias Roux [Univ Paris-Sud]

Post-Doctoral Fellows

Benedikt Becker [Inria, from Jun 2018]
Claudio Belo Da Silva Lourenco [Univ Paris-Sud, from Sep 2018]
Frank Florian Steinberg [Inria, from Oct 2018]

Administrative Assistant

Katia Evrat [Inria]

2. Overall Objectives

2.1. Overall Objectives

The general objective of the Toccata project is to promote formal specification and computer-assisted proof in the development of software that requires high assurance in terms of safety and correctness with respect to the intended behavior of the software.

2.1.1. Context

The importance of software in critical systems increased a lot in the last decade. Critical software appears in various application domains like transportation (e.g., aviation, railway), communication (e.g., smartphones), banking, etc. The number of tasks performed by software is quickly increasing, together with the number of lines of code involved. Given the need of high assurance of safety in the functional behavior of such applications, the need for automated (i.e., computer-assisted) methods and techniques to bring guarantee of safety became a major challenge. In the past and at present, the most widely used approach to check safety of software is to apply heavy test campaigns. These campaigns take a large part of the costs of software development, yet they cannot ensure that all the bugs are caught.

Generally speaking, software verification approaches pursue three goals: (1) verification should be sound, in the sense that no bugs should be missed, (2) verification should not produce false alarms, or as few as possible (3) it should be as automated as possible. Reaching all three goals at the same time is a challenge. A large class of approaches emphasizes goals (2) and (3): testing, run-time verification, symbolic execution, model checking, etc. Static analysis, such as abstract interpretation, emphasizes goals (1) and (3). Deductive verification emphasizes (1) and (2). The Toccata project is mainly interested in exploring the deductive verification approach, although we also consider the others in some cases.

In the past decade, there has been significant progress made in the domain of deductive program verification. They are emphasized by some success stories of application of these techniques on industrial-scale software. For example, the *Atelier B* system was used to develop part of the embedded software of the Paris metro line 14 [46] and other railroad-related systems; a formally proved C compiler was developed using the Coq proof assistant [112]; Microsoft's hypervisor for highly secure virtualization was verified using VCC [85] and the Z3 prover [133]; the L4-verified project developed a formally verified micro-kernel with high security guarantees, using analysis tools on top of the Isabelle/HOL proof assistant [108]. Another sign of recent progress is the emergence of deductive verification competitions (e.g., VerifyThis [2], VScomp [99]).

Finally, recent trends in the industrial practice for development of critical software is to require more and more guarantees of safety, e.g., the upcoming DO-178C standard for developing avionics software adds to the former DO-178B the use of formal models and formal methods. It also emphasizes the need for certification of the analysis tools involved in the process.

2.1.2. Deductive verification

There are two main families of approaches for deductive verification. Methods in the first family build on top of mathematical proof assistants (e.g., Coq, Isabelle) in which both the model and the program are encoded; the proof that the program meets its specification is typically conducted in an interactive way using the underlying proof construction engine. Methods from the second family proceed by the design of standalone tools taking as input a program in a particular programming language (e.g., C, Java) specified with a dedicated annotation language (e.g., ACSL [45], JML [68]) and automatically producing a set of mathematical formulas (the *verification conditions*) which are typically proved using automatic provers (e.g., Z3, Alt-Ergo [48], CVC3 [44], CVC4).

The first family of approaches usually offers a higher level of assurance than the second, but also demands more work to perform the proofs (because of their interactive nature) and makes them less easy to adopt by industry. Moreover, they do not allow to directly analyze a program written in a mainstream programming language like Java or C. The second kind of approaches has benefited in the past years from the tremendous progress made in SAT and SMT solving techniques, allowing more impact on industrial practices, but suffers from a lower level of trust: in all parts of the proof chain (the model of the input programming language, the VC generator, the back-end automatic prover), potential errors may appear, compromising the guarantee offered. Moreover, while these approaches are applied to mainstream languages, they usually support only a subset of their features.

3. Research Program

3.1. Introduction

In the former ProVal project, we have been working on the design of methods and tools for deductive verification of programs. One of our original skills was the ability to conduct proofs by using automatic provers and proof assistants at the same time, depending on the difficulty of the program, and specifically the difficulty of each particular verification condition. We thus believe that we are in a good position to propose a bridge between the two families of approaches of deductive verification presented above. Establishing this bridge is one of the goals of the Toccata project: we want to provide methods and tools for deductive program verification that can offer both a high amount of proof automation and a high guarantee of validity. Toward this objective, a new axis of research was proposed: the development of *certified* analysis tools that are themselves formally proved correct.

The reader should be aware that the word “certified” in this scientific programme means “verified by a formal specification and a formal proof that the program meets this specification”. This differs from the standard meaning of “certified” in an industrial context where it means a conformance to some rigorous process and/or norm. We believe this is the right term to use, as it was used for the *Certified Compiler* project [112], the new conference series *Certified Programs and Proofs*, and more generally the important topics of *proof certificates*.

In industrial applications, numerical calculations are very common (e.g. control software in transportation). Typically they involve floating-point numbers. Some of the members of Toccata have an internationally recognized expertise on deductive program verification involving floating-point computations. Our past work includes a new approach for proving behavioral properties of numerical C programs using Frama-C/Jessie [42], various examples of applications of that approach [65], the use of the Gappa solver for proving numerical algorithms [132], an approach to take architectures and compilers into account when dealing with floating-point programs [66], [123]. We also contributed to the Handbook of Floating-Point Arithmetic [122]. A representative case study is the analysis and the proof of both the method error and the rounding error of a numerical analysis program solving the one-dimension acoustic wave equation [3] [56]. Our experience led us to a conclusion that verification of numerical programs can benefit a lot from combining automatic and interactive theorem proving [59], [65]. Certification of numerical programs is the other main axis of Toccata.

Our scientific programme in structured into four objectives:

1. deductive program verification;
2. automated reasoning;
3. formalization and certification of languages, tools and systems;
4. proof of numerical programs.

We detail these objectives below.

3.2. Deductive Program Verification

Permanent researchers: A. Charguéraud, S. Conchon, J.-C. Filliâtre, C. Marché, G. Melquiond, A. Paskevich

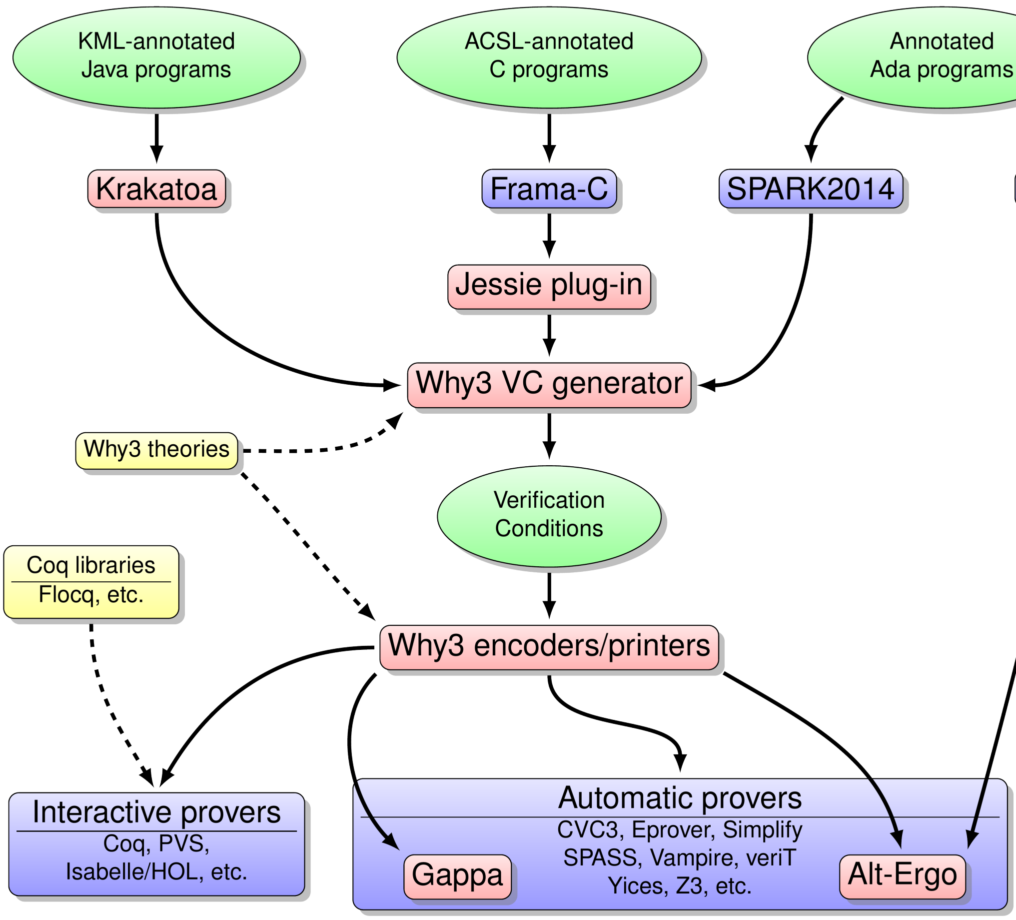


Figure 1. The Why3 ecosystem

3.2.1. The Why3 Ecosystem

This ecosystem is central in our work; it is displayed on Figure 1. The boxes in red background correspond to the tools we develop in the Toccata team.

- The initial design of Why3 was presented in 2012 [51], [98]. In the past years, the main improvements concern the specification language (such as support for higher-order logic functions [72]) and the support for provers. Several new interactive provers are now supported: PVS 6 (used at NASA), Isabelle2014 (planned to be used in the context of Ada program via Spark), and Mathematica. We also added support for new automated provers: CVC4, Metitarski, Metis, Beagle, Princess, and Yices2. More technical improvements are the design of a Coq tactic to call provers via Why3 from Coq, and the design of a proof session mechanism [50]. Why3 was presented during several invited talks [97], [96], [93], [94].
- At the level of the C front-end of Why3 (via Frama-C), we have proposed an approach to add a notion of refinement on C programs [131], and an approach to reason about pointer programs with a standard logic, via *separation predicates* [49]
- The Ada front-end of Why3 has mainly been developed during the past three years, leading to the release of SPARK2014 [107] (http://www.spark-2014.org/)
- In collaboration with J. Almeida, M. Barbosa, J. Pinto, and B. Vieira (University do Minho, Braga, Portugal), J.-C. Filliâtre has developed a method for certifying programs involving cryptographic methods. It uses Why as an intermediate language [41].
- With M. Pereira and S. Melo de Sousa (Universidade da Beira Interior, Covilhã, Portugal), J.-C. Filliâtre has developed an environment for proving ARM assembly code. It uses Why3 as an intermediate VC generator. It was presented at the Inforum conference [126] (best student paper).

3.2.2. Concurrent Programming

- S. Conchon and A. Mebsout, in collaboration with F. Zaïdi (VALS team, LRI), A. Goel and S. Krstić (Strategic Cad Labs, INTEL) have proposed a new model-checking approach for verifying safety properties of array-based systems. This is a syntactically restricted class of parametrized transition systems with states represented as arrays indexed by an arbitrary number of processes. Cache coherence protocols and mutual exclusion algorithms are typical examples of such systems. It was first presented at CAV 2012 [5] and detailed further [83]. It was applied to the verification of programs with fences [79]. The core algorithm has been extended with a mechanism for inferring invariants. This new algorithm, called BRAB, is able to automatically infer invariants strong enough to prove industrial cache coherence protocols. BRAB computes over-approximations of backward reachable states that are checked to be unreachable in a finite instance of the system. These approximations (candidate invariants) are then model-checked together with the original safety properties. Completeness of the approach is ensured by a mechanism for backtracking on spurious traces introduced by too coarse approximations [80], [118].
- In the context of the ERC DeepSea project [0], A. Charguéraud and his co-authors have developed a unifying semantics for various different paradigms of parallel computing (fork-join, async-finish, and futures), and published a conference paper describing this work [40]. Besides, A. Charguéraud and his co-authors have polished their previous work on granularity control for parallel algorithms using user-provided complexity functions, and produced a journal article [39].

3.2.3. Case Studies

- To provide an easy access to the case studies that we develop using Why3 and its front-ends, we have published a *gallery of verified programs* on our web page http://toccata.lri.fr/gallery/. Part of these examples are the solutions to the competitions VerifyThis 2011 [67], VerifyThis 2012 [2], and the competition VScomp 2011 [99].

[0] Arthur Charguéraud is involved 40% of his time in the ERC DeepSea project, which is hosted at Inria Paris Rocquencourt (team Gallium).

- Other case studies that led to publications are the design of a library of data-structures based on AVLs [71], the verification a two-lines C program (solving the N-queens puzzle) using Why3 [95], and the verification of Koda and Ruskey's algorithm [100].
- A. Charguéraud, with F. Pottier (Inria Paris), extended their formalization of the correctness and asympotic complexity of the classic Union Find data structure, which features the bound expressed in terms of the inverse Ackermann function [38]. The proof, conducted using CFML extended with time credits, was refined using a slightly more complex potential function, allowing to derive a simpler and richer interface for the data structure [70].

For other case studies, see also sections of numerical programs and formalization of languages and tools.

3.2.4. Project-team Positioning

Several research groups in the world develop their own approaches, techniques, and tools for deductive verification. With respect to all these related approaches and tools, our originality is our will to use more sophisticated specification languages (with inductive definitions, higher-order features and such) and the ability to use a large set of various theorem provers, including the use of interactive theorem proving to deal with complex functional properties.

- The RiSE team [0] at Microsoft Research Redmond, USA, partly in collaboration with team "programming methodology" team [0] at ETH Zurich develop tools that are closely related to ours: Boogie and Dafny are direct competitors of Why3, VCC is a direct competitor of Frama-C/Jessie.
- The KeY project [0] (several teams, mainly at Karlsruhe and Darmstadt, Germany, and Göteborg, Sweden) develops the KeY tool for Java program verification [37], based on dynamic logic, and has several industrial users. They use a specific modal logic (dynamic logic) for modeling programs, whereas we use standard logic, so as to be able to use off-the-shelf automated provers.
- The "software engineering" group at Augsburg, Germany, develops the KIV system [0], which was created more than 20 years ago (1992) and is still well maintained and efficient. It provides a semi-interactive proof environment based on algebraic-style specifications, and is able to deal with several kinds of imperative style programs. They have a significant industrial impact.
- The VeriFast system [0] aims at verifying C programs specified in Separation Logic. It is developed at the Katholic University at Leuven, Belgium. We do not usually use separation logic (so as to use off-the-shelf provers) but alternative approaches (e.g. static memory separation analysis).
- The Mobius Program Verification Environment [0] is a joint effort for the verification of Java source annotated with JML, combining static analysis and runtime checking. The tool ESC/Java2 [0] is a VC generator similar to Krakatoa, that builds on top of Boogie. It is developed by a community leaded by University of Copenhagen, Denmark. Again, our specificity with respect to them is the consideration of more complex specification languages and interactive theorem proving.
- The Lab for Automated Reasoning and Analysis [0] at EPFL, develop methods and tools for verification of Java (Jahob) and Scala (Leon) programs. They share with us the will and the ability to use several provers at the same time.
- The TLA environment [0], developed by Microsoft Research and the Inria team Veridis, aims at the verification of concurrent programs using mathematical specifications, model checking, and interactive or automated theorem proving.
- The F* project [0], developed by Microsoft Research and the Inria Prosecco team, aims at providing a rich environment for developing programs and proving them.

[0] http://research.microsoft.com/en-us/groups/rise/default.aspx
[0] http://www.pm.inf.ethz.ch/
[0] http://www.key-project.org/
[0] http://www.isse.uni-augsburg.de/en/software/kiv/
[0] http://people.cs.kuleuven.be/~bart.jacobs/verifast/
[0] http://kindsoftware.com/products/opensource/Mobius/
[0] http://kindsoftware.com/products/opensource/ESCJava2/
[0] http://lara.epfl.ch/w/
[0] http://research.microsoft.com/en-us/um/people/lamport/tla/tla.html
[0] http://research.microsoft.com/en-us/projects/fstar/

The KeY and KIV environments mentioned above are partly based on interactive theorem provers. There are other approaches on top of general-purpose proof assistants for proving programs that are not purely functional:

- The Ynot project [0] is a Coq library for writing imperative programs specified in separation logic. It was developed at Harvard University, until the end of the project in 2010. Ynot had similar goals as CFML, although Ynot requires programs to be written in monadic style inside Coq, whereas CFML applies directly on programs written in OCaml syntax, translating them into logical formulae.
- Front-ends to Isabelle were developed to deal with simple sequential imperative programs [130] or C programs [125]. The L4-verified project [108] is built on top of Isabelle.

3.3. Automated Reasoning

Permanent researchers: S. Conchon, G. Melquiond, A. Paskevich

3.3.1. Generalities on Automated Reasoning

- J. C. Blanchette and A. Paskevich have designed an extension to the TPTP TFF (Typed First-order Form) format of theorem proving problems to support rank-1 polymorphic types (also known as ML-style parametric polymorphism) [47]. This extension, named TFF1, has been incorporated in the TPTP standard.
- S. Conchon defended his *habilitation à diriger des recherches* in December 2012. The memoir [76] provides a useful survey of the scientific work of the past 10 years, around the SMT solving techniques, that led to the tools Alt-Ergo and Cubicle as they are nowadays.

3.3.2. Quantifiers and Triggers

- C. Dross, J. Kanig, S. Conchon, and A. Paskevich have proposed a generic framework for adding a decision procedure for a theory or a combination of theories to an SMT prover. This mechanism is based on the notion of instantiation patterns, or *triggers*, which restrict instantiation of universal premises and can effectively prevent a combinatorial explosion. A user provides an axiomatization with triggers, along with a proof of completeness and termination in the proposed framework, and obtains in return a sound, complete and terminating solver for his theory. A prototype implementation was realized on top of Alt-Ergo. As a case study, a feature-rich axiomatization of doubly-linked lists was proved complete and terminating [88]. C. Dross defended her PhD thesis in April 2014 [89]. The main results of the thesis are: (1) a formal semantics of the notion of *triggers* typically used to control quantifier instantiation in SMT solvers, (2) a general setting to show how a first-order axiomatization with triggers can be proved correct, complete, and terminating, and (3) an extended DPLL(T) algorithm to integrate a first-order axiomatization with triggers as a decision procedure for the theory it defines. Significant case studies were conducted on examples coming from SPARK programs, and on the benchmarks on B set theory constructed within the BWare project.

3.3.3. Reasoning Modulo Theories

- S. Conchon, É. Contejean and M. Iguernelala have presented a modular extension of ground AC-completion for deciding formulas in the combination of the theory of equality with user-defined AC symbols, uninterpreted symbols and an arbitrary signature-disjoint Shostak theory X [78]. This work extends the results presented in [77] by showing that a simple preprocessing step allows to get rid of a full AC-compatible reduction ordering, and to simply use a partial multiset extension of a *non-necessarily AC-compatible* ordering.
- S. Conchon, M. Iguernelala, and A. Mebsout have designed a collaborative framework for reasoning modulo simple properties of non-linear arithmetic [82]. This framework has been implemented in the Alt-Ergo SMT solver.

[0] http://ynot.cs.harvard.edu/

- S. Conchon, G. Melquiond and C. Roux have described a dedicated procedure for a theory of floating-point numbers which allows reasoning on approximation errors. This procedure is based on the approach of the Gappa tool: it performs saturation of consequences of the axioms, in order to refine bounds on expressions. In addition to the original approach, bounds are further refined by a constraint solver for linear arithmetic [84]. This procedure has been implemented in Alt-Ergo.
- In collaboration with A. Mahboubi (Inria project-team Typical), and G. Melquiond, the group involved in the development of Alt-Ergo have implemented and proved the correctness of a novel decision procedure for quantifier-free linear integer arithmetic [1]. This algorithm tries to bridge the gap between projection and branching/cutting methods: it interleaves an exhaustive search for a model with bounds inference. These bounds are computed provided an oracle capable of finding constant positive linear combinations of affine forms. An efficient oracle based on the Simplex procedure has been designed. This algorithm is proved sound, complete, and terminating and is implemented in Alt-Ergo.
- Most of the results above are detailed in M. Iguernelala's PhD thesis [105].

3.3.4. Applications

- We have been quite successful in the application of Alt-Ergo to industrial development: qualification by Airbus France, integration of Alt-Ergo into the Spark Pro toolset.
- In the context of the BWare project, aiming at using Why3 and Alt-Ergo for discharging proof obligations generated by Atelier B, we made progress into several directions. The method of translation of B proof obligations into Why3 goals was first presented at ABZ'2012 [121]. Then, new drivers have been designed for Why3, in order to use new back-end provers Zenon modulo and iProver modulo. A notion of rewrite rule was introduced into Why3, and a transformation for simplifying goals before sending them to back-end provers was designed. Intermediate results obtained so far in the project were presented both at the French conference AFADL [87] and at ABZ'2014 [86].

  On the side of Alt-Ergo, recent developments have been made to efficiently discharge proof obligations generated by Atelier B. This includes a new plugin architecture to facilitate experiments with different SAT engines, new heuristics to handle quantified formulas, and important modifications in its internal data structures to boost performances of core decision procedures. Benchmarks realized on more than 10,000 proof obligations generated from industrial B projects show significant improvements [81].
- Hybrid automatons interleave continuous behaviors (described by differential equations) with discrete transitions. D. Ishii and G. Melquiond have worked on an automated procedure for verifying safety properties (that is, global invariants) of such systems [106].

3.3.5. Project-team Positioning

Automated Theorem Proving is a large community, but several sub-groups can be identified:

- The SMT-LIB community gathers people interested in reasoning modulo theories. In this community, only a minority of participants are interested in supporting first-order quantifiers at the same time as theories. SMT solvers that support quantifiers are Z3 (Microsoft Research Redmond, USA), CVC3 and its successor CVC4 [0].
- The TPTP community gathers people interested in first-order theorem proving.
- Other Inria teams develop provers: veriT by team Veridis, and Psyche by team Parsifal.
- Other groups develop provers dedicated to very specific cases, such as Metitarski [0] at Cambridge, UK, which aims at proving formulas on real numbers, in particular involving special functions such as log or exp. The goal is somewhat similar to our CoqInterval library, *cf* objective 4.

[0] http://cvc4.cs.stanford.edu/web/
[0] http://www.cl.cam.ac.uk/~lp15/papers/Arith/

It should be noticed that a large number of provers mentioned above are connected to Why3 as back-ends.

3.4. Formalization and Certification of Languages, Tools and Systems

Permanent researchers: S. Boldo, A. Charguéraud, C. Marché, G. Melquiond, C. Paulin

3.4.1. Real Numbers, Real Analysis, Probabilities

- S. Boldo, C. Lelay, and G. Melquiond have worked on the Coquelicot library, designed to be a user-friendly Coq library about real analysis [62], [63]. An easier way of writing formulas and theorem statements is achieved by relying on total functions in place of dependent types for limits, derivatives, integrals, power series, and so on. To help with the proof process, the library comes with a comprehensive set of theorems and some automation. We have exercised the library on several use cases: on an exam at university entry level [110], for the definitions and properties of Bessel functions [109], and for the solution of the one-dimensional wave equation [111]. We have also conducted a survey on the formalization of real arithmetic and real analysis in various proof systems [64].
- Watermarking techniques are used to help identify copies of publicly released information. They consist in applying a slight and secret modification to the data before its release, in a way that should remain recognizable even in (reasonably) modified copies of the data. Using the Coq ALEA library, which formalizes probability theory and probabilistic programs, D. Baelde together with P. Courtieu, D. Gross-Amblard from Rennes and C. Paulin have established new results about the robustness of watermarking schemes against arbitrary attackers [43]. The technique for proving robustness is adapted from methods commonly used for cryptographic protocols and our work illustrates the strengths and particularities of the ALEA style of reasoning about probabilistic programs.

3.4.2. Formalization of Languages, Semantics

- P. Herms, together with C. Marché and B. Monate (CEA List), has developed a certified VC generator, using Coq. The program for VC calculus and its specifications are both written in Coq, but the code is crafted so that it can be extracted automatically into a stand-alone executable. It is also designed in a way that allows the use of arbitrary first-order theorem provers to discharge the generated obligations [104]. On top of this generic VC generator, P. Herms developed a certified VC generator for C source code annotated using ACSL. This work is the main result of his PhD thesis [103].
- A. Tafat and C. Marché have developed a certified VC generator using Why3 [114], [115]. The challenge was to formalize the operational semantics of an imperative language, and a corresponding weakest precondition calculus, without the possibility to use Coq advanced features such as dependent types or higher-order functions. The classical issues with local bindings, names and substitutions were solved by identifying appropriate lemmas. It was shown that Why3 can offer a significantly higher amount of proof automation compared to Coq.
- A. Charguéraud, together with Alan Schmitt (Inria Rennes) and Thomas Wood (Imperial College), has developed an interactive debugger for JavaScript. The interface, accessible as a webpage in a browser, allows to execute a given JavaScript program, following step by step the formal specification of JavaScript developped in prior work on *JsCert* [52]. Concretely, the tool acts as a double-debugger: one can visualize both the state of the interpreted program and the state of the interpreter program. This tool is intended for the JavaScript committee, VM developpers, and other experts in JavaScript semantics.
- M. Clochard, C. Marché, and A. Paskevich have developed a general setting for developing programs involving binders, using Why3. This approach was successfully validated on two case studies: a verified implementation of untyped lambda-calculus and a verified tableaux-based theorem prover [75].

- M. Clochard, J.-C. Filliâtre, C. Marché, and A. Paskevich have developed a case study on the formalization of semantics of programming languages using Why3 [72]. This case study aims at illustrating recent improvements of Why3 regarding the support for higher-order logic features in the input logic of Why3, and how these are encoded into first-order logic, so that goals can be discharged by automated provers. This case study also illustrates how reasoning by induction can be done without need for interactive proofs, via the use of *lemma functions*.
- M. Clochard and L. Gondelman have developed a formalization of a simple compiler in Why3 [73]. It compiles a simple imperative language into assembler instructions for a stack machine. This case study was inspired by a similar example developed using Coq and interactive theorem proving. The aim is to improve significantly the degree of automation in the proofs. This is achieved by the formalization of a Hoare logic and a Weakest Precondition Calculus on assembly programs, so that the correctness of compilation is seen as a formal specification of the assembly instructions generated.

3.4.3. Project-team Positioning

The objective of formalizing languages and algorithms is very general, and it is pursued by several Inria teams. One common trait is the use of the Coq proof assistant for this purpose: Pi.r2 (development of Coq itself and its meta-theory), Gallium (semantics and compilers of programming languages), Marelle (formalization of mathematics), SpecFun (real arithmetic), Celtique (formalization of static analyzers).

Other environments for the formalization of languages include

- ACL2 system [0]: an environment for writing programs with formal specifications in first-order logic based on a Lisp engine. The proofs are conducted using a prover based on the Boyer-Moore approach. It is a rather old system but still actively maintained and powerful, developed at University of Texas at Austin. It has a strong industrial impact.
- Isabelle environment [0]: both a proof assistant and an environment for developing pure applicative programs. It is developed jointly at University of Cambridge, UK, Technische Universität München, Germany, and to some extent by the VALS team at LRI, Université Paris-Sud. It features highly automated tactics based on ATP systems (the Sledgehammer tool).
- The team "Trustworthy Systems" at NICTA in Australia [0] aims at developing highly trustable software applications. They developed a formally verified micro-kernel called seL4 [108], using a home-made layer to deal with C programs on top of the Isabelle prover.
- The PVS system [0] is an environment for both programming and proving (purely applicative) programs. It is developed at the Computer Science Laboratory of SRI international, California, USA. A major user of PVS is the team LFM [0] at NASA Langley, USA, for the certification of programs related to air traffic control.

In the Toccata team, we do not see these alternative environments as competitors, even though, for historical reasons, we are mainly using Coq. Indeed both Isabelle and PVS are available as back-ends of Why3.

3.5. Proof of Numerical Programs

Permanent researchers: S. Boldo, C. Marché, G. Melquiond

- Linked with objective 1 (Deductive Program Verification), the methodology for proving numerical C programs has been presented by S. Boldo in her habilitation [54] and as invited speaker [55]. An application is the formal verification of a numerical analysis program. S. Boldo, J.-C. Filliâtre, and G. Melquiond, with F. Clément and P. Weis (POMDAPI team, Inria Paris - Rocquencourt), and M. Mayero (LIPN), completed the formal proof of the second-order centered finite-difference scheme for the one-dimensional acoustic wave [57][3].

[0] http://www.cs.utexas.edu/~moore/acl2/
[0] http://isabelle.in.tum.de/
[0] http://ssrg.nicta.com.au/projects/TS/
[0] http://pvs.csl.sri.com/
[0] http://shemesh.larc.nasa.gov/fm/fm-main-team.html

- Several challenging floating-point algorithms have been studied and proved. This includes an algorithm by Kahan for computing the area of a triangle: S. Boldo proved an improvement of its error bound and new investigations in case of underflow [53]. This includes investigations about quaternions. They should be of norm 1, but due to the round-off errors, a drift of this norm is observed over time. C. Marché determined a bound on this drift and formally proved it correct [9]. P. Roux formally verified an algorithm for checking that a matrix is semi-definite positive [129]. The challenge here is that testing semi-definiteness involves algebraic number computations, yet it needs to be implemented using only approximate floating-point operations.
- Because of compiler optimizations (or bugs), the floating-point semantics of a program might change once compiled, thus invalidating any property proved on the source code. We have investigated two ways to circumvent this issue, depending on whether the compiler is a black box. When it is, T. Nguyen has proposed to analyze the assembly code it generates and to verify it is correct [124]. On the contrary, S. Boldo and G. Melquiond (in collaboration with J.-H. Jourdan and X. Leroy) have added support for floating-point arithmetic to the CompCert compiler and formally proved that none of the transformations the compiler applies modify the floating-point semantics of the program [61], [60].
- Linked with objectives 2 (Automated Reasoning) and 3 (Formalization and Certification of Languages, Tools and Systems), G. Melquiond has implemented an efficient Coq library for floating-point arithmetic and proved its correctness in terms of operations on real numbers [119]. It serves as a basis for an interval arithmetic on which Taylor models have been formalized. É. Martin-Dorel and G. Melquiond have integrated these models into CoqInterval [10]. This Coq library is dedicated to automatically proving the approximation properties that occur when formally verifying the implementation of mathematical libraries (libm).
- Double rounding occurs when the target precision of a floating-point computation is narrower than the working precision. In some situations, this phenomenon incurs a loss of accuracy. P. Roux has formally studied when it is innocuous for basic arithmetic operations [129]. É. Martin-Dorel and G. Melquiond (in collaboration with J.-M. Muller) have formally studied how it impacts algorithms used for error-free transformations [117]. These works were based on the Flocq formalization of floating-point arithmetic for Coq.
- By combining multi-precision arithmetic, interval arithmetic, and massively-parallel computations, G. Melquiond (in collaboration with G. Nowak and P. Zimmermann) has computed enough digits of the Masser-Gramain constant to invalidate a 30-year old conjecture about its closed form [120].

3.5.1. Project-team Positioning

This objective deals both with formal verification and floating-point arithmetic, which is quite uncommon. Therefore our competitors/peers are few. We may only cite the works by J. Duracz and M. Konečný, Aston University in Birmingham, UK.

The Inria team AriC (Grenoble - Rhône-Alpes) is closer to our research interests, but they are lacking manpower on the formal proof side; we have numerous collaborations with them. The Inria team Caramel (Nancy - Grand Est) also shares some research interests with us, though fewer; again, they do not work on the formal aspect of the verification; we have some occasional collaborations with them.

There are many formalization efforts from chip manufacturers, such as AMD (using the ACL2 proof assistant) and Intel (using the Forte proof assistants) but the algorithms they consider are quite different from the ones we study. The works on the topic of floating-point arithmetic from J. Harrison at Intel using HOL Light are really close to our research interests, but they seem to be discontinued.

A few deductive program verification teams are willing to extend their tools toward floating-point programs. This includes the KeY project and SPARK. We have an ongoing collaboration with the latter, in the context of the ProofInUSe project.

Deductive verification is not the only way to prove programs. Abstract interpretation is widely used, and several teams are interested in floating-point arithmetic. This includes the Inria team Antique (Paris - Rocquencourt) and a CEA List team, who have respectively developed the Astrée and Fluctuat tools. This approach targets a different class of numerical algorithms than the ones we are interested in.

Other people, especially from the SMT community (*cf* objective 2), are also interested in automatically proving formulas about floating-point numbers, notably at Oxford University. They are mainly focusing on pure floating-point arithmetic though and do not consider them as approximation of real numbers.

Finally, it can be noted that numerous teams are working on the verification of numerical programs, but assuming the computations are real rather than floating-point ones. This is out of the scope of this objective.

4. Application Domains

4.1. Domain 1

The application domains we target involve safety-critical software, that is where a high-level guarantee of soundness of functional execution of the software is wanted. Currently our industrial collaborations mainly belong to the domain of transportation, including aeronautics, railroad, space flight, automotive.

Verification of C programs, Alt-Ergo at Airbus Transportation is the domain considered in the context of the ANR U3CAT project, led by CEA, in partnership with Airbus France, Dassault Aviation, Sagem Défense et Sécurité. It included proof of C programs via Frama-C/Jessie/Why, proof of floating-point programs [116], the use of the Alt-Ergo prover via CAVEAT tool (CEA) or Frama-C/WP. Within this context, we contributed to a qualification process of Alt-Ergo with Airbus industry: the technical documents (functional specifications and benchmark suite) have been accepted by Airbus, and these documents were submitted by Airbus to the certification authorities (DO-178B standard) in 2012. This action is continued in the new project Soprano.

Certified compilation, certified static analyzers Aeronautics is the main target of the Verasco project, led by Verimag, on the development of certified static analyzers, in partnership with Airbus. This is a follow-up of the transfer of the CompCert certified compiler (Inria team Gallium) to which we contributed to the support of floating-point computations [61].

Transfer to the community of Ada development The former FUI project Hi-Lite, led by Adacore company, introduced the use of Why3 and Alt-Ergo as back-end to SPARK2014, an environment for verification of Ada programs. This is applied to the domain of aerospace (Thales, EADS Astrium). At the very beginning of that project, Alt-Ergo was added in the Spark Pro toolset (predecessor of SPARK2014), developed by Altran-Praxis: Alt-Ergo can be used by customers as an alternate prover for automatically proving verification conditions. Its usage is described in the new edition of the Spark book [0] (Chapter “Advanced proof tools”). This action is continued in the new joint laboratory ProofInUse. A recent paper [69] provides an extensive list of applications of SPARK, a major one being the British air control management *iFacts*.

Transfer to the community of Atelier B In the current ANR project BWare, we investigate the use of Why3 and Alt-Ergo as an alternative back-end for checking proof obligations generated by *Atelier B*, whose main applications are railroad-related software [0], a collaboration with Mitsubishi Electric R&D Centre Europe (Rennes) (joint publication [121]) and ClearSy (Aix-en-Provence).

SMT-based Model-Checking: Cubicle S. Conchon (with A. Mebsout and F. Zaidi from VALS team at LRI) has a long-term collaboration with S. Krstic and A. Goel (Intel Strategic Cad Labs in Hillsboro, OR, USA) that aims in the development of the SMT-based model checker Cubicle (http://cubicle.lri.fr/) based on Alt-Ergo [118][5]. It is particularly targeted to the verification of concurrent programs and protocols.

[0] http://www.altran-praxis.com/book/
[0] http://www.methode-b.com/

5. Highlights of the Year

5.1. Highlights of the Year

J.-C. Filliâtre served as judge at the ICPC regional programming contests SWERC 2017 and 2018. These two editions were organized in Paris and gathered each year 80 teams of three students from universities and schools from South-West Europe. https://swerc.eu/

The 2nd edition of the Handbook of Floating-Point arithmetic was published [28]

5.1.1. Awards

R. Rieu-Helft received the "Student Gold Medal" award, and J.-C. Filliâtre the "Best challenge submitted" award, at the *VerifyThis@ETAPS2018 verification competition*http://www.pm.inf.ethz.ch/research/verifythis/Prizes.html

6. New Software and Platforms

6.1. Alt-Ergo

Automated theorem prover for software verification
KEYWORDS: Software Verification - Automated theorem proving
FUNCTIONAL DESCRIPTION: Alt-Ergo is an automatic solver of formulas based on SMT technology. It is especially designed to prove mathematical formulas generated by program verification tools, such as Frama-C for C programs, or SPARK for Ada code. Initially developed in Toccata research team, Alt-Ergo's distribution and support are provided by OCamlPro since September 2013.

RELEASE FUNCTIONAL DESCRIPTION: the "SAT solving" part can now be delegated to an external plugin, new experimental SAT solver based on mini-SAT, provided as a plugin. This solver is, in general, more efficient on ground problems, heuristics simplification in the default SAT solver and in the matching (instantiation) module, re-implementation of internal literals representation, improvement of theories combination architecture, rewriting some parts of the formulas module, bugfixes in records and numbers modules, new option "-no-Ematching" to perform matching without equality reasoning (i.e. without considering "equivalence classes"). This option is very useful for benchmarks coming from Atelier-B, two new experimental options: "-save-used-context" and "-replay-used-context". When the goal is proved valid, the first option allows to save the names of useful axioms into a ".used" file. The second one is used to replay the proof using only the axioms listed in the corresponding ".used" file. Note that the replay may fail because of the absence of necessary ground terms generated by useless axioms (that are not included in .used file) during the initial run.

- Participants: Alain Mebsout, Évelyne Contejean, Mohamed Iguernelala, Stéphane Lescuyer and Sylvain Conchon
- Partner: OCamlPro
- Contact: Sylvain Conchon
- URL: http://alt-ergo.lri.fr

6.2. CoqInterval

Interval package for Coq
KEYWORDS: Interval arithmetic - Coq
FUNCTIONAL DESCRIPTION: CoqInterval is a library for the proof assistant Coq.

It provides several tactics for proving theorems on enclosures of real-valued expressions. The proofs are performed by an interval kernel which relies on a computable formalization of floating-point arithmetic in Coq.

The Marelle team developed a formalization of rigorous polynomial approximation using Taylor models in Coq. In 2014, this library has been included in CoqInterval.

- Participants: Assia Mahboubi, Érik Martin-Dorel, Guillaume Melquiond, Jean-Michel Muller, Laurence Rideau, Laurent Théry, Micaela Mayero, Mioara Joldes, Nicolas Brisebarre and Thomas Sibut-Pinote
- Contact: Guillaume Melquiond
- Publications: Proving bounds on real-valued functions with computations - Floating-point arithmetic in the Coq system - Proving Tight Bounds on Univariate Expressions with Elementary Functions in Coq - Formally Verified Approximations of Definite Integrals - Formally Verified Approximations of Definite Integrals
- URL: http://coq-interval.gforge.inria.fr/

6.3. Coquelicot

The Coquelicot library for real analysis in Coq
KEYWORDS: Coq - Real analysis
FUNCTIONAL DESCRIPTION: Coquelicot is library aimed for supporting real analysis in the Coq proof assistant. It is designed with three principles in mind. The first is the user-friendliness, achieved by implementing methods of automation, but also by avoiding dependent types in order to ease the stating and readability of theorems. This latter part was achieved by defining total function for basic operators, such as limits or integrals. The second principle is the comprehensiveness of the library. By experimenting on several applications, we ensured that the available theorems are enough to cover most cases. We also wanted to be able to extend our library towards more generic settings, such as complex analysis or Euclidean spaces. The third principle is for the Coquelicot library to be a conservative extension of the Coq standard library, so that it can be easily combined with existing developments based on the standard library.

- Participants: Catherine Lelay, Guillaume Melquiond and Sylvie Boldo
- Contact: Sylvie Boldo
- URL: http://coquelicot.saclay.inria.fr/

6.4. Cubicle

The Cubicle model checker modulo theories
KEYWORDS: Model Checking - Software Verification
FUNCTIONAL DESCRIPTION: Cubicle is an open source model checker for verifying safety properties of array-based systems, which corresponds to a syntactically restricted class of parametrized transition systems with states represented as arrays indexed by an arbitrary number of processes. Cache coherence protocols and mutual exclusion algorithms are typical examples of such systems.

- Participants: Alain Mebsout and Sylvain Conchon
- Contact: Sylvain Conchon
- URL: http://cubicle.lri.fr/

6.5. Flocq

The Flocq library for formalizing floating-point arithmetic in Coq
KEYWORDS: Floating-point - Arithmetic code - Coq
FUNCTIONAL DESCRIPTION: The Flocq library for the Coq proof assistant is a comprehensive formalization of floating-point arithmetic: core definitions, axiomatic and computational rounding operations, high-level properties. It provides a framework for developers to formally verify numerical applications.

Flocq is currently used by the CompCert verified compiler to support floating-point computations.

- Participants: Guillaume Melquiond, Pierre Roux and Sylvie Boldo
- Contact: Sylvie Boldo
- Publications: Flocq: A Unified Library for Proving Floating-point Algorithms in Coq - A Formally-Verified C Compiler Supporting Floating-Point Arithmetic - Verified Compilation of Floating-Point Computations - Innocuous Double Rounding of Basic Arithmetic Operations - Formal Proofs of Rounding Error Bounds - Computer Arithmetic and Formal Proofs
- URL: http://flocq.gforge.inria.fr/

6.6. Gappa

The Gappa tool for automated proofs of arithmetic properties
KEYWORDS: Floating-point - Arithmetic code - Software Verification - Constraint solving
FUNCTIONAL DESCRIPTION: Gappa is a tool intended to help formally verifying numerical programs dealing with floating-point or fixed-point arithmetic. It has been used to write robust floating-point filters for CGAL and it is used to verify elementary functions in CRlibm. While Gappa is intended to be used directly, it can also act as a backend prover for the Why3 software verification plateform or as an automatic tactic for the Coq proof assistant.

- Participant: Guillaume Melquiond
- Contact: Guillaume Melquiond
- Publications: Generating formally certified bounds on values and round-off errors - Formal certification of arithmetic filters for geometric predicates - Assisted verification of elementary functions - From interval arithmetic to program verification - Formally Certified Floating-Point Filters For Homogeneous Geometric Predicates - Combining Coq and Gappa for Certifying Floating-Point Programs - Handbook of Floating-Point Arithmetic - Certifying the floating-point implementation of an elementary function using Gappa - Automations for verifying floating-point algorithms - Automating the verification of floating-point algorithms - Computer Arithmetic and Formal Proofs
- URL: http://gappa.gforge.inria.fr/

6.7. Why3

The Why3 environment for deductive verification
KEYWORDS: Formal methods - Trusted software - Software Verification - Deductive program verification
FUNCTIONAL DESCRIPTION: Why3 is an environment for deductive program verification. It provides a rich language for specification and programming, called WhyML, and relies on external theorem provers, both automated and interactive, to discharge verification conditions. Why3 comes with a standard library of logical theories (integer and real arithmetic, Boolean operations, sets and maps, etc.) and basic programming data structures (arrays, queues, hash tables, etc.). A user can write WhyML programs directly and get correct-by-construction OCaml programs through an automated extraction mechanism. WhyML is also used as an intermediate language for the verification of C, Java, or Ada programs.

- Participants: Andriy Paskevych, Claude Marché, François Bobot, Guillaume Melquiond, Jean-Christophe Filliâtre, Levs Gondelmans and Martin Clochard
- Partners: CNRS - Université Paris-Sud
- Contact: Claude Marché
- URL: http://why3.lri.fr/

6.8. Coq

The Coq Proof Assistant
KEYWORDS: Proof - Certification - Formalisation

SCIENTIFIC DESCRIPTION: Coq is an interactive proof assistant based on the Calculus of (Co-)Inductive Constructions, extended with universe polymorphism. This type theory features inductive and co-inductive families, an impredicative sort and a hierarchy of predicative universes, making it a very expressive logic. The calculus allows to formalize both general mathematics and computer programs, ranging from theories of finite structures to abstract algebra and categories to programming language metatheory and compiler verification. Coq is organised as a (relatively small) kernel including efficient conversion tests on which are built a set of higher-level layers: a powerful proof engine and unification algorithm, various tactics/decision procedures, a transactional document model and, at the very top an IDE.
FUNCTIONAL DESCRIPTION: Coq provides both a dependently-typed functional programming language and a logical formalism, which, altogether, support the formalisation of mathematical theories and the specification and certification of properties of programs. Coq also provides a large and extensible set of automatic or semi-automatic proof methods. Coq's programs are extractible to OCaml, Haskell, Scheme, ...

RELEASE FUNCTIONAL DESCRIPTION: Coq version 8.8.2 contains the result of refinements and stabilization of features and deprecations, cleanups of the internals of the system along with a few new features.

Summary of changes:

Kernel: fix a subject reduction failure due to allowing fixpoints on non-recursive values (#407), by Matthieu Sozeau. Handling of evars in the VM (#935) by Pierre-Marie Pédrot.

Notations: many improvements on recursive notations and support for destructuring patterns in the syntax of notations by Hugo Herbelin.

Proof language: tacticals for profiling, timing and checking success or failure of tactics by Jason Gross. The focusing bracket { supports single-numbered goal selectors, e.g. 2:{, (#6551) by Théo Zimmermann.

Vernacular: cleanup of definition commands (#6653) by Vincent Laporte and more uniform handling of the Local flag (#1049), by Maxime Dénès. Experimental Show Extraction command (#6926) by Pierre Letouzey. Coercion now accepts Prop or Type as a source (#6480) by Arthur Charguéraud. Export modifier for options allowing to export the option to modules that Import and not only Require a module (#6923), by Pierre-Marie Pédrot.

Universes: many user-level and API level enhancements: qualified naming and printing, variance annotations for cumulative inductive types, more general constraints and enhancements of the minimization heuristics, interaction with modules by Gaëtan Gilbert, Pierre-Marie Pédrot and Matthieu Sozeau.

Library: Decimal Numbers library (#6599) by Pierre Letouzey and various small improvements.

Documentation: a large community effort resulted in the migration of the reference manual to the Sphinx documentation tool. The new documentation infrastructure (based on Sphinx) is by Clément Pit-Claudel. The migration was coordinated by Maxime Dénès and Paul Steckler, with some help of Théo Zimmermann during the final integration phase. The 14 people who ported the manual are Calvin Beck, Heiko Becker, Yves Bertot, Maxime Dénès, Richard Ford, Pierre Letouzey, Assia Mahboubi, Clément Pit-Claudel, Laurence Rideau, Matthieu Sozeau, Paul Steckler, Enrico Tassi, Laurent Théry, Nikita Zyuzin.

Tools: experimental -mangle-names option to coqtop/coqc for linting proof scripts (#6582), by Jasper Hugunin. Main changes:

Critical soundness bugs were fixed between versions 8.8.0 and 8.8.2, and a PDF version of the reference manual was made available. The Windows installer also includes many more external packages that can be individually selected for installation.

On the implementation side, the dev/doc/changes.md file documents the numerous changes to the implementation and improvements of interfaces. The file provides guidelines on porting a plugin to the new version.

More information can be found in the CHANGES file. Feedback and bug reports are extremely welcome.

Distribution Installers for Windows 32 bits (i686), Windows 64 bits (x8_64) and macOS are available. They come bundled with CoqIDE. Windows binaries now include the Bignums library.

Complete sources of the files installed by the Windows installers are made available, to comply with license requirements.
NEWS OF THE YEAR: Version 8.8.0 was released in April 2018 and version 8.8.2 in September 2018. This is the third release of Coq developed on a time-based development cycle. Its development spanned 6 months from the release of Coq 8.7 and was based on a public road-map. It attracted many external contributions. Code reviews and continuous integration testing were systematically used before integration of new features, with an important focus given to compatibility and performance issues.

The main advances in this version are cleanups and fixes in the many different components of the system, ranging from low level kernel fixes to advances in the support of notations and tacticals for selecting goals. A large community effort was made to move the documentation to the Sphinx format, providing a more accessible online ressource to users.

- Participants: Abhishek Anand, C. J. Bell, Yves Bertot, Frédéric Besson, Tej Chajed, Pierre Courtieu, Maxime Denes, Julien Forest, Emilio Jesús Gallego Arias, Gaëtan Gilbert, Benjamin Grégoire, Jason Gross, Hugo Herbelin, Ralf Jung, Matej Kosik, Sam Pablo Kuper, Xavier Leroy, Pierre Letouzey, Assia Mahboubi, Cyprien Mangin, Érik Martin-Dorel, Olivier Marty, Guillaume Melquiond, Pierre-Marie Pédrot, Benjamin C. Pierce, Lars Rasmusson, Yann Régis-Gianas, Lionel Rieg, Valentin Robert, Thomas Sibut-Pinote, Michael Soegtrop, Matthieu Sozeau, Arnaud Spiwack, Paul Steckler, George Stelle, Pierre-Yves Strub, Enrico Tassi, Hendrik Tews, Laurent Théry, Amin Timany, Vadim Zaliva and Théo Zimmermann
- Partners: CNRS - Université Paris-Sud - ENS Lyon - Université Paris-Diderot
- Contact: Matthieu Sozeau
- Publication: The Coq Proof Assistant, version 8.8.0
- URL: http://coq.inria.fr/

7. New Results

7.1. Deductive Verification

Synthetic topology in HoTT for probabilistic programming. F. Faissole and B. Spitters have developed a mathematical formalism based on synthetic topology and homotopy type theory to interpret probabilistic algorithms. They suggest to use proof assistants to prove such programs [91] [92]. They also have formalized synthetic topology in the Coq proof assistant using the HoTT library. It consists of a theory of lower reals, valuations and lower integrals. All the results are constructive. They apply their results to interpret probabilistic programs using a monadic approach [23].

A Toolchain to Produce Correct-by-Construction OCaml Programs In the context of the research project Vocal, J.-C. Filliâtre, A. Paskevich, and M. Pereira, together with L. Gondelman (postdoc in January 2017) and S. Melo de Sousa (visiting Associate Professor from UBI, Portugal, in Sep/Oct 2017), designed and implemented a toolchain for the verification of OCaml code using Why3 [33]. In this framework, the user provides a formal specification within comments embedded in the OCaml interface file together with an implementation in Why3. Two tools automatically translate the former to a Why3 specification and the latter to an OCaml code. One the refinement proof is completed on the Why3 side, the overall diagram commutes, ensuring the soundness of the OCaml code.

Ghost monitors M. Clochard, C. Marché, and A. Paskevich designed a new approach to deductive program verification based on auxiliary programs called *ghost monitors*. This technique is useful when the syntactic structure of the target program is not well suited for verification, for example, when an essentially recursive algorithm is implemented in an iterative fashion. The approach consists in implementing, specifying, and verifying an auxiliary program that monitors the execution of the target program, in such a way that the correctness of the monitor entails the correctness of the target.

This technique is also applicable when one wants to establish relational properties between two target programs written in different languages and having different syntactic structure [32] [29].

This approach is based on an earlier variant proposed in M. Clochard's PhD thesis [11]. The ghost monitor maintains the necessary data and invariants to facilitate the proof, it can be implemented and verified in any suitable framework, which does not have to be related to the language of the target programs. M. Clochard introduced one such framework, with an original extension that allows one to specify and prove fine-grained properties about infinite behaviors of target programs. The proof of correctness of this approach relies on a particular flavor of transfinite games. This proof is formalized and verified using the Why3 tool (http://toccata.lri.fr/gallery/hoare_logic_and_games.en.html).

Extracting Why3 programs to C programs. R. Rieu-Helft, C. Marché, and G. Melquiond devised a simple memory model for representing C-like pointers in the Why3 system. This makes it possible to translate a small fragment of Why3 verified programs into idiomatic C code [26]. This extraction mechanism was used to turn a verified Why3 library of arbitrary-precision integer arithmetic into a C library that can be substituted to part of the GNU Multi-Precision (GMP) library [128].

Verification of highly imperative OCaml programs with Why3 J.-C. Filliâtre, M. Pereira, and S. Melo de Sousa proposed a new methodology for proving highly imperative OCaml programs with Why3. For a given OCaml program, a specific memory model is built and one checks a Why3 program that operates on it. Once the proof is complete, they use Why3's extraction mechanism to translate its programs to OCaml, while replacing the operations on the memory model with the corresponding operations on mutable types of OCaml. This method is evaluated on several examples that manipulate linked lists and mutable graphs [24].

Verification of Parameterized Concurrent Programs on Weak Memory Models Modern multiprocessors and microprocesseurs implement weak or relaxed memory models, in which the apparent order of memory operation does not follow the sequential consistency (SC) proposed by Leslie Lamport. Any concurrent program running on such architecture and designed with an SC model in mind may exhibit new behaviors during its execution, some of which may potentially be incorrect. For instance, a mutual exclusion algorithm, correct under an interleaving semantics, may no longer guarantee mutual exclusion when implemented on a weaker architecture. Reasoning about the semantics of such programs is a difficult task. Moreover, most concurrent algorithms are designed for an arbitrary number of processes. D. Declerck [12] proposed an approach to ensure the correctness of such concurrent algorithms, regardless of the number of processes involved. It relies on the Model Checking Modulo Theories (MCMT) framework, developed by Ghilardi and Ranise, which allows for the verification of safety properties of parameterized concurrent programs, that is to say, programs involving an arbitrary number of processes. This technology is extended with a theory for reasoning about weak memory models. The result is an extension of the Cubicle model checker called Cubicle-W, which allows the verification of safety properties of parameterized transition systems running under a weak memory model similar to TSO.

Counterexample Generation S. Dailler and C. Marché worked on extensions and improvements of the counterexample generation feature of Why3, used in particular by the SPARK front-end for Ada [102] [101]. When the logic goal generated for a given verification condition is not shown unsatisfiable by an SMT solvers, some solver can propose a model. By carefully reverting the transformation chain (from an input program through the VC generator and the various translation steps to solvers), this model is turned into a potential counterexample that the user can exploit to analyze why its original code is not proved. The extension consists in a deep analysis of the complete model generated by the solver, so as to extract more information and produce better counterexamples. A journal paper giving the details of the whole process was published [14]

Alias Control for SPARK Program Verification G.-A. Jaloyan and A. Paskevich, together with C. Dross, M. Maalej, and Y. Moy made a proposal for introduction of pointers to the SPARK language, based on permission-driven static alias analysis method inspired by Rust's borrow-checker and affine types [35]. By ensuring that at any point of execution any writable value can only be accessed through a single name, it is possible to apply the standard rules of Hoare logic (or weakest precondition calculus) to verify programs with pointers. The proposed framework was implemented in the GNAT Ada compiler and the SPARK toolset.

7.2. Automated Reasoning

A Why3 Framework for Reflection Proofs and its Application to GMP's Algorithms Earlier works using Why3 showed that automatically verifying the algorithms of the arbitrary-precision integer library GMP exceeds the current capabilities of automatic solvers. To complete this verification, numerous cut indications had to be supplied by the user, slowing the project to a crawl. G. Melquiond and R. Rieu-Helf extended Why3 with a framework for proofs by reflection, with minimal impact on the trusted computing base. This framework makes it easy to write dedicated decision procedures that make full use of Why3's imperative features and are formally verified. This approach opens the way to efficiently tackling the further verification of GMP's algorithms [20], [27].

Expressive and extensible automated reasoning tactics for Coq Proof assistants based on Type Theory, such as Coq, allow implementing effective automatic tactics based on computational reasoning (e.g. `lia` for linear integer arithmetic, or `ring` for ring theory). Unfortunately, these are usually limited to one particular domain. In contrast, SMTCoq is a modular and extensible tool, using external provers, which generalizes these computational approaches to combine multiple theories. It relies on a high-level interface, which offers a greater expressiveness, at the cost of more complex automation. Q. Garchery, in collaboration with C. Keller and V .Blot, designed two improvements to increase expressiveness of SMTCoq without impeding its modularity and its efficiency: the first adds some support for universally quantified hypotheses, while the second generalizes the support for integer arithmetic to the different representations of natural numbers and integers in Coq. This work will be presented in the next JFLA [30]

Non-linear Arithmetic Reasoning for Control-Command Software State-of-the-art (semi-)decision procedures for non-linear real arithmetic address polynomial inequalities by mean of symbolic methods, such as quantifier elimination, or numerical approaches such as interval arithmetic. Although (some of) these methods offer nice completeness properties, their high complexity remains a limit, despite the impressive efficiency of modern implementations. This appears to be an obstacle to the use of SMT solvers when verifying, for instance, functional properties of control-command programs. Using off-the-shelf convex optimization solvers is known to constitute an appealing alternative. However, these solvers only deliver approximate solutions, which means they do not readily provide the soundness expected for applications such as software verification. S. Conchon, together with P. Roux and M. Iguernelala [21], investigated a-posteriori validation methods and their integration in the SMT framework. Although their early prototype, implemented in the Alt-Ergo SMT solver, often does not prove competitive with state of the art solvers, it already gives some interesting results, particularly on control-command programs.

Lightweight Interactive Proving for Automated Program Verification Deductive verification approach allows establishing the strongest possible formal guarantees on critical software. The downside is the cost in terms of human effort required to design adequate formal specifications and to successfully discharge the required proof obligations. To popularize deductive verification in an industrial software development environment, it is essential to provide means to progressively transition from simple and automated approaches to deductive verification. The SPARK environment, for development of critical software written in Ada, goes towards this goal by providing automated tools for formally proving that some code fulfills the requirements expressed in Ada contracts.

In a program verifier that makes use of automatic provers to discharge the proof obligations, a need for some additional user interaction with proof tasks shows up: either to help analyzing the reason of a proof failure or, ultimately, to discharge the verification conditions that are out-of-reach of state-of-the-art automatic provers. Adding interactive proof features in SPARK appears to be complicated by the fact that the proof toolchain makes use of the independent, intermediate verification tool Why3, which is generic enough to accept multiple front-ends for different input languages. S. Dailler, C. Marché and Y. Moy proposed an approach to extend Why3 with interactive proof features and also with a generic client-server infrastructure allowing integration of proof interaction into an external, front-end graphical user interface such as the one of SPARK. This was presented at the F-IDE symposium [18].

7.3. Certification of Algorithms, Languages, Tools and Systems

Formalization and closedness of finite dimensional subspaces. F. Faissole formalized a theory of finite dimensional subspaces of Hilbert spaces in order to apply the Lax-Milgram Theorem on such subspaces. He had to prove, in the Coq proof assistant, that finite dimensional subspaces of Hilbert spaces are closed in the context of general topology using filters [90]. He also formalized both finite dimensional modules and finite dimensional subspaces of modules. He compared the two formalizations and showed a complementarity between them. He proved that the product of two finite dimensional modules is a finite dimensional module [22].

Analysis of explicit Runge-Kutta methods Numerical integration schemes are mandatory to understand complex behaviors of dynamical systems described by ordinary differential equations. Implementation of these numerical methods involve floating-point computations and propagation of round-off errors. In the spirit of [58], S. Boldo, F. Faissole and A. Chapoutot developed a fine-grained analysis of round-off errors in explicit Runge-Kutta integration methods, taking into account exceptional behaviors, such as underflow and overflow [31].

Verified numerical approximations of improper definite integrals. The CoqInterval library provides some tactics for computing and formally verifying numerical approximations of real-valued expressions inside the Coq system. In particular, it is able to compute reliable bounds on proper definite integrals [113]. A. Mahboubi, G. Melquiond, and T. Sibut-Pinote extended these algorithms to also cover some improper integrals, e.g., those with an unbounded integration domain [15]. This makes CoqInterval one of the very few tools able to produce reliable results for improper integrals, be they formally verified or not.

Case study: algorithms for matrix multiplication. M. Clochard, L. Gondelman and M. Pereira worked on a case study about matrix multiplication. Two variants for the multiplication of matrices are proved: a naive version using three nested loops and Strassen's algorithm. To formally specify the two multiplication algorithms, they developed a new Why3 theory of matrices, and they applied a reflection methodology to conduct some of the proofs. A first version of this work was presented at the VSTTE Conference in 2016 [74]. An extended version that considers arbitrary rectangular matrices instead of square ones is published in the Journal of Automated Reasoning [13]. The development is available in Toccata's gallery http://toccata.lri.fr/gallery/verifythis_2016_matrix_multiplication.en.html.

Digital Filters Digital filters are small iterative algorithms, used as basic bricks in signal processing (filters) and control theory (controllers). D. Gallois-Wong, S. Boldo and T. Hilaire formally proved in Coq some error analysis theorems about digital filters, namely the Worst-Case Peak Gain theorem and the existence of a filter characterizing the difference between the exact filter and the implemented one. Moreover, as the digital signal processing literature provides many equivalent algorithms, called realizations, they formally defined and proved the equivalence of several realizations (Direct Forms and State-Space) [19]. Another Coq development dedicated the a realization called SIF (Specialized Implicit Form) has been done, in order to encompass all the other realizations up to the order of computation, which is very important in finite precision [25].

7.4. Floating-Point and Numerical Programs

Correct Average of Decimal Floating-Point Numbers Some modern processors include decimal floating-point units, with a conforming implementation of the IEEE-754 2008 standard. Unfortunately, many algorithms from the computer arithmetic literature are not correct anymore when computations are done in radix 10. This is in particular the case for the computation of the average of two floating-point numbers. S. Boldo, F. Faissole and V. Tourneur developed a new radix-10 algorithm that computes the correctly-rounded average, with a Coq formal proof of its correctness, that takes gradual underflow into account [17].

Optimal Inverse Projection of Floating-Point Addition In a setting where we have intervals for the values of floating-point variables x, a, and b, we are interested in improving these intervals when the floating-point equality $x \oplus a = b$ holds. This problem is common in constraint propagation, and called the inverse projection of the addition. It also appears in abstract interpretation for the analysis of programs containing IEEE 754 operations. D. Gallois-Wong, S. Boldo and P. Cuoq proposed floating-point theorems that provide optimal bounds for all the intervals. Fast loop-free algorithms compute these optimal bounds using only floating-point computations at the target precision [34].

Handbook of Floating-point Arithmetic Initially published in 2010, the *Handbook of Floating-Point Arithmetic* has been heavily updated. G. Melquiond contributed to the second edition [28].

Error analysis of finite precision digital filters and controllers The effort to provide accurate and reliable error analysis of fixed-point implementations of Signal Processing and Control algorithms was continued (see also the formalization effort above). A. Volkova, M. Istoan, F. de Dinechin and T. Hilaire (Citi Lyon, INSA Lyon) created an automatic code generator for FPGAs and dedicated roundoff analysis in order to minimize the bit-widths used for the intern computations while guaranteeing a bound on the output error [16]. The global workflow for the rigorous design of reliable Fixed-Point filters has been studied by A. Volkova, T. Hilaire and C. Lauter and submitted to a journal [36] : it concerns the rigorous determination of the Most Significant Bit of each variable, to guaranty that no overflow will ever occur, also taking into account the roundoff error propagation.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. ProofInUse Joint Laboratory

Participants: Claude Marché [contact], Jean-Christophe Filliâtre, Andrei Paskevich, Guillaume Melquiond, Sylvain Dailler.

The objective of ProofInUse is to provide verification tools, based on mathematical proof, to industry users. These tools are aimed at replacing or complementing the existing test activities, whilst reducing costs.

This laboratory is a joint effort of the Inria project-team Toccata, the AdaCore company which provides development tools for the Ada programming language, and the TrustInSoft company which provides static analysis tools for the C and C++ programming language.

The objective of ProofInUse is thus to significantly increase the capabilities and performances of verification environments proposed by these two companies. It aims at integration of verification techniques at the state-of-the-art of academic research, via the generic environment Why3 for deductive program verification developed by Toccata.

This joint laboratory is a follow-up of the former "LabCom ProofInUse" between Toccata and AdaCore, funded by the ANR programme "Laboratoires communs", from April 2014 to March 2017 http://www.spark-2014.org/proofinuse.

The SME AdaCore is a software publisher specializing in providing software development tools for critical systems. A previous successful collaboration between Toccata and AdaCore enabled *Why3* technology to be put into the heart of the AdaCore-developed SPARK technology.

The SME TrustInSoft is a company whose speciality is the verification of critical software, written in the C or C++ languages. It is interested in integrating the novelties of ProofInUse in its own environment TIS Analyzer.

8.2. Bilateral Grants with Industry

8.2.1. CIFRE contract with TrustInSoft company

Participants: Guillaume Melquiond [contact], Raphaël Rieu-Helft.

Jointly with the thesis of R. Rieu-Helft, supervised in collaboration with the TrustInSoft company, we established a 3-year bilateral collaboration contract, that started in October 2017. The aim is to design methods that make it possible to design an arbitrary-precision integer library that, while competitive with the state-of-the-art library GMP, is formally verified. Not only are GMP's algorithm especially intricate from an arithmetic point of view, but numerous tricks were also used to optimize them. We are using the Why3 programming language to implement the algorithms, we are developing reflection-based procedures to verify them, and we finally extract them as a C library that is binary-compatible with GMP [20] [26].

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. ELEFFAN

Participant: Sylvie Boldo [contact].

ELEFFAN is a Digicosme project funding the PhD of F. Faissole. S. Boldo is the principal investigator. It began in 2016 for three years. https://project.inria.fr/eleffan/

The ELEFFAN project aims at formally proving rounding error bounds of numerical schemes.

Partners: ENSTA Paristech (A. Chapoutot)

9.1.2. MILC

Participant: Sylvie Boldo [contact].

MILC is a DIM-RFSI project. It is a one-year project (2018–2019) that aims at formalizing measure theory and Lebesgue integral in the Coq proof assistant. https://lipn.univ-paris13.fr/MILC/

Partners: Université Paris 13 (M. Mayero, PI), Inria Paris, Inria Saclay

9.2. National Initiatives

9.2.1. ANR CoLiS

Participants: Claude Marché [contact], Andrei Paskevich.

The CoLiS research project is funded by the programme "Société de l'information et de la communication" of the ANR, for a period of 60 months, starting on October 1st, 2015. http://colis.irif.univ-paris-diderot.fr/

The project aims at developing formal analysis and verification techniques and tools for scripts. These scripts are written in the POSIX or bash shell language. Our objective is to produce, at the end of the project, formal methods and tools allowing to analyze, test, and validate scripts. For this, the project will develop techniques and tools based on deductive verification and tree transducers stemming from the domain of XML documents.

Partners: Université Paris-Diderot, IRIF laboratory (formerly PPS & LIAFA), coordinator; Inria Lille, team LINKS

9.2.2. ANR Vocal

Participants: Jean-Christophe Filliâtre [contact], Andrei Paskevich.

The Vocal research project is funded by the programme "Société de l'information et de la communication" of the ANR, for a period of 60 months, starting on October 1st, 2015. https://vocal.lri.fr/

The goal of the Vocal project is to develop the first formally verified library of efficient general-purpose data structures and algorithms. It targets the OCaml programming language, which allows for fairly efficient code and offers a simple programming model that eases reasoning about programs. The library will be readily available to implementers of safety-critical OCaml programs, such as Coq, Astrée, or Frama-C. It will provide the essential building blocks needed to significantly decrease the cost of developing safe software. The project intends to combine the strengths of three verification tools, namely Coq, Why3, and CFML. It will use Coq to obtain a common mathematical foundation for program specifications, as well as to verify purely functional components. It will use Why3 to verify a broad range of imperative programs with a high degree of proof automation. Finally, it will use CFML for formal reasoning about effectful higher-order functions and data structures making use of pointers and sharing.

Partners: team Gallium (Inria Paris-Rocquencourt), team DCS (Verimag), TrustInSoft, and OCamlPro.

9.2.3. ANR FastRelax

Participants: Sylvie Boldo [contact], Guillaume Melquiond.

This is a research project funded by the programme "Ingénierie Numérique & Sécurité" of the ANR. It is funded for a period of 48 months and it has started on October 1st, 2014. http://fastrelax.gforge.inria.fr/

Our aim is to develop computer-aided proofs of numerical values, with certified and reasonably tight error bounds, without sacrificing efficiency. Applications to zero-finding, numerical quadrature or global optimization can all benefit from using our results as building blocks. We expect our work to initiate a "fast and reliable" trend in the symbolic-numeric community. This will be achieved by developing interactions between our fields, designing and implementing prototype libraries and applying our results to concrete problems originating in optimal control theory.

Partners: team ARIC (Inria Grenoble Rhône-Alpes), team MARELLE (Inria Sophia Antipolis - Méditerranée), team SPECFUN (Inria Saclay - Île-de-France), Université Paris 6, and LAAS (Toulouse).

9.2.4. ANR Soprano

Participants: Sylvain Conchon [contact], Guillaume Melquiond.

The Soprano research project is funded by the programme "Sciences et technologies logicielles" of the ANR, for a period of 42 months, starting on October 1st, 2014. http://soprano-project.fr/

The SOPRANO project aims at preparing the next generation of verification-oriented solvers by gathering experts from academia and industry. We will design a new framework for the cooperation of solvers, focused on model generation and borrowing principles from SMT (current standard) and CP (well-known in optimization). Our main scientific and technical objectives are the following. The first objective is to design a new collaboration framework for solvers, centered around synthesis rather than satisfiability and allowing cooperation beyond that of Nelson-Oppen while still providing minimal interfaces with theoretical guarantees. The second objective is to design new decision procedures for industry-relevant and hard-to-solve theories. The third objective is to implement these results in a new open-source platform. The fourth objective is to ensure industrial-adequacy of the techniques and tools developed through periodical evaluations from the industrial partners.

Partners: team DIVERSE (Inria Rennes - Bretagne Atlantique), Adacore, CEA List, Université Paris-Sud, and OCamlPro.

9.2.5. FUI LCHIP

Participant: Sylvain Conchon [contact].

LCHIP (Low Cost High Integrity Platform) is aimed at easing the development of safety critical applications (up to SIL4) by providing: (i) a complete IDE able to automatically generate and prove bounded complexity software (ii) a low cost, safe execution platform. The full support of DSLs and third party code generators will enable a seamless deployment into existing development cycles. LCHIP gathers scientific results obtained during the last 20 years in formal methods, proof, refinement, code generation, etc. as well as a unique return of experience on safety critical systems design. http://www.clearsy.com/en/2016/10/4260/

Partners: 2 technology providers (ClearSy, OcamlPro), in charge of building the architecture of the platform; 3 labs (IFSTTAR, LIP6, LRI), to improve LCHIP IDE features; 2 large companies (SNCF, RATP), representing public ordering parties, to check compliance with standard and industrial railway use-case.

The project lead by ClearSy has started in April 2016 and lasts 3 years. It is funded by BpiFrance as well as French regions.

9.2.6. ANR PARDI

Participant: Sylvain Conchon [contact].

Verification of PARameterized DIstributed systems. A parameterized system specification is a specification for a whole class of systems, parameterized by the number of entities and the properties of the interaction, such as the communication model (synchronous/asynchronous, order of delivery of message, application ordering) or the fault model (crash failure, message loss). To assist and automate verification without parameter instantiation, PARDI uses two complementary approaches. First, a fully automatic model checker modulo theories is considered. Then, to go beyond the intrinsic limits of parameterized model checking, the project advocates a collaborative approach between proof assistant and model checker. http://pardi.enseeiht.fr/

The proof lead by Toulouse INP/IRIT started in 2016 and lasts for 4 years. Partners: Université Pierre et Marie Curie (LIP6), Université Paris-Sud (LRI), Inria Nancy (team VERIDIS)

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

9.3.1.1. EMC2

Participant: Sylvie Boldo [contact].

A new ERC Synergy Grant 2018 project, called Extreme-scale Mathematically-based Computational Chemistry (EMC2) has just been accepted. The PIs are É. Cancès, L. Grigori, Y. Maday and J.-P. Piquemal. S. Boldo is part of the work package 3: validation and certification of molecular simulation results. https://www.sorbonne-universite.fr/newsroom/actualites/erc-synergy-grant-2018

9.3.2. Collaborations in European Programs, Except FP7 & H2020

Program: COST (European Cooperation in Science and Technology).

Project acronym: EUTypes https://eutypes.cs.ru.nl/

Project title: The European research network on types for programming and verification

Duration: 2015-2019

Coordinator: Herman Geuvers, Radboud University Nijmegen, The Netherlands

Other partners: 36 members countries, see http://www.cost.eu/COST_Actions/ca/CA15123?parties

Abstract: Types are pervasive in programming and information technology. A type defines a formal interface between software components, allowing the automatic verification of their connections, and greatly enhancing the robustness and reliability of computations and communications. In rich dependent type theories, the full functional specification of a program can be expressed as a type. Type systems have rapidly evolved over the past years, becoming more sophisticated, capturing new aspects of the behaviour of programs and the dynamics of their execution.

This COST Action will give a strong impetus to research on type theory and its many applications in computer science, by promoting (1) the synergy between theoretical computer scientists, logicians and mathematicians to develop new foundations for type theory, for example as based on the recent development of "homotopy type theory", (2) the joint development of type theoretic tools as proof assistants and integrated programming environments, (3) the study of dependent types for programming and its deployment in software development, (4) the study of dependent types for verification and its deployment in software analysis and verification. The action will also tie together these different areas and promote cross-fertilisation.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

S. Boldo, president of the 29th “Journées Francophones des Langages Applicatifs” (JFLA 2018)

J.-C. Filliâtre, scientific chair and co-organizer of EJCP (École Jeunes Chercheurs en Programmation du GDR GPL) at Lyon on June 25–29, 2018. 5 days / 8 lectures / 25 participants. https://ejcp2018.sciencesconf.org/

D. Gallois-Wong, co-chair of the Doctoral Programme of the 11th Conference on Intelligent Computer Mathematics (CICM 2018).

10.1.1.2. Member of the Organizing Committees

G. Melquiond, organizer of the 10th “Rencontres Arithmétiques du GDR Informatique-Mathématique” (RAIM 2018)

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

S. Boldo, program chair of the 29th “Journées Francophones des Langages Applicatifs” (JFLA 2018).

S. Boldo, program co-chair of the 26th IEEE Symposium on Computer Arithmetic (ARITH 2019), Kyoto, Japan.

10.1.2.2. Member of the Conference Program Committees

S. Boldo, PC of the 25th IEEE Symposium on Computer Arithmetic (ARITH 2018)

S. Boldo, PC of the 7th ACM SIGPLAN Conference on Certified Programs and Proofs (CPP 2018)

S. Boldo, PC of the Tenth NASA Formal Methods Symposium (NFM 2018)

S. Boldo, PC of the Eleventh NASA Formal Methods Symposium (NFM 2019)

J.-C. Filliâtre, PC of the 18th International Workshop on Automated Verification of Critical Systems (AVOCS 2018)

J.-C. Filliâtre, PC of the 10th International Conference on Interactive Theorem Proving (ITP 2019)

J.-C. Filliâtre, PC of the European Symposium on Programming (ESOP 2020)

J.-C. Filliâtre, PC of the Symposium on Languages, Applications and Technologies (SLATE 2018)

G. Melquiond, PC of the 26th IEEE Symposium on Computer Arithmetic (ARITH 2019)

G. Melquiond, PC of the 10th International Conference on Interactive Theorem Proving (ITP 2019)

10.1.2.3. Reviewer

The members of the Toccata team have reviewed papers for numerous international conferences.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

G. Melquiond, member of the editorial board of *Reliable Computing*.

J.-C. Filliâtre, member of the editorial board of *Journal of Functional Programming*.

10.1.3.2. Reviewer - Reviewing Activities

The members of the Toccata team have reviewed numerous papers for numerous international journals.

10.1.4. Invited Talks

J.-C. Filliâtre, invited speaker at the 8th International Conference on Interactive Theorem Proving (ITP 2018).

J.-C. Filliâtre, invited speaker at the Formal Integrated Development Environment (F-IDE 2018).

10.1.5. Leadership within the Scientific Community

S. Boldo, elected chair of the ARITH working group of the GDR-IM (a CNRS subgroup of computer science) with J. Detrey (Inria Nancy).

J.-C. Filliâtre, chair of IFIP WG 1.9/2.15 verified Software.

10.1.6. Scientific Expertise

G. Melquiond, member of the scientific commission of Inria-Saclay, in charge of selecting candidates for PhD grants, Post-doc grants, temporary leaves from universities ("délégations").

C. Marché, member of the "Bureau du Comité des Projets" of Inria-Saclay (includes examination of proposals for creation of new Inria project-teams for Saclay research center).

S. Boldo, member of the program committee for selecting postdocs of the maths/computer science program of the Labex mathématique Hadamard.

S. Boldo, member of the national Inria admission committee.

J.-C. Filliâtre, grading the entrance examination at X/ENS (*"option informatique"*).

C. Marché, scientific expert for project evaluation, Dutch Research Council (NWO https://www.nwo.nl/en), The Netherlands, 2018.

C. Marché, scientific expert for project evaluation, National Science Centre (Narodowe Centrum Nauki - NCN http://www.ncn.gov.pl/), Poland, 2018.

C. Marché, scientific expert for promotion of academic staff, Chalmers University of Technology, Sweden, 2018.

S. Boldo, member of a hiring committee for an associate professor position in computer science at University Paris Diderot (IRIF laboratory).

C. Marché, member of DigiCosme committee for research and innovation (selection of projects for working groups, post-doc grants, doctoral missions, invited professors)

10.1.7. Research Administration

G. Melquiond, member of the committee for the monitoring of PhD students (*"commission de suivi doctoral"*).

S Boldo, member of the CLFP (*"commission locale de formation permanente"*).

S. Boldo, member of the CCD, (*"commission consultative des doctorants"*).

S. Boldo will be deputy scientific director (DSA) of Inria Saclay research center from January 1st, 2019

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

J.-C. Filliâtre, *Langages de programmation et compilation*, 25h, L3, École Normale Supérieure, France.

J.-C. Filliâtre, *Les bases de l'algorithmique et de la programmation*, 15h, L3, École Polytechnique, France.

J.-C. Filliâtre, *Compilation*, 18h, M1, École Polytechnique, France.

G. Melquiond, *Programmation C++ avancée*, 12h, M2, Université Paris-Saclay, France.

10.2.2. Supervision

PhD: M. Clochard, "Methods and tools for specification and proof of difficult properties of sequential programs" [11], Université Paris-Saclay & Université Paris-Sud, March 30th 2018, supervised by C. Marché and A. Paskevich.

PhD: D. Declerck, "Verification via Model Checking of Parameterized Concurrent Programs on Weak Memory Models" [12], Université Paris-Saclay & Université Paris-Sud, Sep 24th 2018, supervised by F. Zaïdi (LRI) and S. Conchon.

PhD: M. Pereira, "Tools and Techniques for the Verification of Modular Stateful Code" [127], Université Paris-Saclay & Université Paris-Sud, Dec 10th 2018, supervised by by J.-C. Filliâtre.

PhD in progress: M. Roux, "Model Checking de systèmes paramétrés et temporisés", since Sep. 2015, supervised by Sylvain Conchon.

PhD in progress: A. Coquereau, "[ErgoFast] Amélioration de performances pour le solveur SMT Alt-Ergo : conception d'outils d'analyse, optimisations et structures de données efficaces pour OCaml", since Sep. 2015, supervised by S. Conchon, F. Le Fessant et M. Mauny.

PhD in progress: F. Faissole, "Stabilité(s): liens entre l'arithmétique flottante et l'analyse numérique", since Oct. 2016, supervised by S. Boldo and A. Chapoutot.

PhD in progress: R. Rieu-Helft, "Développement et vérification de bibliothèques d'arithmétique entière en précision arbitraire", since Oct. 2017, supervised by G. Melquiond and P. Cuoq (TrustIn-Soft).

PhD in progress: D. Gallois-Wong, "Vérification formelle et filtres numériques", since Oct. 2017, supervised by S. Boldo and T. Hilaire.

PhD in progress: Q. Garchery, "Certification de la génération et de la transformation d'obligations de preuve", since Oct. 2018, supervised by C. Keller, C. Marché and A. Paskevich.

10.2.3. Juries

C. Marché: examiner of the habilitation thesis of J. Signoles, "From Static Analysis to Runtime Verification with Frama-C and E-ACSL", Université Paris-Sud, July 9th 2018

C. Marché: examiner of the habilitation thesis of N. Kosmatov, "Combinations of Analysis Techniques for Sound and Efficient Software Verification", Université Paris-Sud, Nov 20th 2018

C. Marché: president of the PhD defense of J.-C. Léchenet, "Certified Algorithms for Program Slicing", Université Paris-Saclay, July 19th 2018

C. Marché: reviewer of the PhD defense of C. Laurenço, "Single-assignment Program Verification", Universidad do Minho, Portugal, July 2nd 2018

S. Boldo: reviewer and president of the PhD defense of B. Djalal, "Formalisation en Coq pour la décision de problèmes en géométrie algébrique réelle", Université Côte d'Azur, December 3rd 2018

S. Boldo: reviewer of the PhD of R. Picot, "Amélioration de la fiabilité numérique de codes de calcul industriels", Sorbonne Université, March 27th 2018

S. Boldo: president of the PhD defense of S. Covanov, "Algorithmes de multiplication : complexité bilinéaire et méthodes asymptotiquement rapides", Université de Lorraine, June 5th 2018

S. Boldo: president of the PhD defense of G. Davy, “Génération de codes et d’annotations prouvables d’algorithmes de points intérieurs à destination de systèmes embarqués critiques”, Université de Toulouse, December 6th 2018

J.-C. Filliâtre: licentiate doctorate examination at Chalmers University of Technology, Sweden, August 23, 2018.

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

S. Boldo is the scientific head for Saclay for the MECSI group for networking about computer science popularization inside Inria.

She was also responsible (with M. Quet of the SCM) for the 2018 “Fête de la science” on October 11th 2018. About 260 teenagers were welcomed on 8 activities ranging from unplugged activities with Duplo construction toys to programming, and from applied mathematics to theoretical computer science.

10.3.2. Interventions

S. Boldo animated an activity at the Inria “Fête de la science” on October 11th 2018 the whole day long.

S. Boldo animated an activity and gave talks at the LRI “Fête de la science” on October 12th 2018.

S. Boldo gave a talk during at a *Girls can code* week on August 31st 2018 in Paris.

S. Boldo will give a talk to about 180 teenagers at the Marie Curie high school in Sceaux on February 8th, 2019

J.-C. Filliâtre gave a talk at *Mathematical Summer in Paris* on July 16, 2018.

J.-C. Filliâtre gave a talk *Parcours d’un informaticien* at the seminar *“Info Pour Tous”* (high school and undergraduate students). Video on YouTube. http://seminairespourtous.ens.fr/ipt

S. Dailler and C. Marché gave a demonstration of the SPARK environment, at the DigiHall Day (May 22 2018 https://www.irt-systemx.fr/evenements/digihall-2018/). DigiHall is a cluster of digital technologies of Paris-Saclay. More than 800 industrial and institutional decision-makers and academic counterparts took part in this first-of-its-kind event.

C. Marché presented the joint laboratory ProofInUse at the LabCom Colloquium (Maison de la Chimie, Paris, Sep. 27 2018 http://ptolemee.com/colloque-labcom/index.html) organized by ANR, with participation of numerous actors from both academia and industry.

10.3.3. Internal action

S. Boldo demonstrated popularization by an unplugged activity to all the new Inria staff at the welcome days on June 7th 2018

S. Boldo animated an unplugged activity to the AER service (team assistants) on July 3rd 2018

S. Boldo trained colleagues on unplugged activities for the “Fête de la science” (5 sessions of about 1h30)

S. Dailler and C. Marché gave a presentation of the joint laboratory ProofInUse, together with a demonstration of the SPARK environment, at the Software Day of the DigiCosme Labex (Saclay, June 7 2018 https://digicosme.lri.fr/tiki-read_article.php?articleId=256)

10.3.4. Creation of media or tools for science outreach

S. Boldo is supervising the popularization mission of C. Patte (M3DISIM team) in order to create a new popularization activity for teenagers in 2019.

C. Marché, main contributor of the site for the Why3 tool inside the Inria Saclay Virtual Showroom. Includes a short video introduction of Why3 for beginners using the TryWhy3 Web interface http://why3.lri.fr/try/

11. Bibliography

Major publications by the team in recent years

[1] F. BOBOT, S. CONCHON, É. CONTEJEAN, M. IGUERNELALA, A. MAHBOUBI, A. MEBSOUT, G. MELQUIOND.*A Simplex-Based Extension of Fourier-Motzkin for Solving Linear Integer Arithmetic*, in "IJCAR 2012: Proceedings of the 6th International Joint Conference on Automated Reasoning", Manchester, UK, B. GRAMLICH, D. MILLER, U. SATTLER (editors), Lecture Notes in Computer Science, Springer, June 2012, vol. 7364, p. 67–81, http://hal.inria.fr/hal-00687640

[2] F. BOBOT, J.-C. FILLIÂTRE, C. MARCHÉ, A. PASKEVICH.*Let's Verify This with Why3*, in "International Journal on Software Tools for Technology Transfer (STTT)", 2015, vol. 17, n^{o} 6, p. 709–727, http://hal.inria.fr/hal-00967132/en

[3] S. BOLDO, F. CLÉMENT, J.-C. FILLIÂTRE, M. MAYERO, G. MELQUIOND, P. WEIS.*Wave Equation Numerical Resolution: a Comprehensive Mechanized Proof of a C Program*, in "Journal of Automated Reasoning", April 2013, vol. 50, n^{o} 4, p. 423–456, http://hal.inria.fr/hal-00649240/en/

[4] S. BOLDO, G. MELQUIOND.*Computer Arithmetic and Formal Proofs: Verifying Floating-point Algorithms with the Coq System*, ISTE Press - Elsevier, December 2017, https://hal.inria.fr/hal-01632617

[5] S. CONCHON, A. GOEL, S. KRSTIĆ, A. MEBSOUT, F. ZAÏDI.*Cubicle: A Parallel SMT-based Model Checker for Parameterized Systems*, in "CAV 2012: Proceedings of the 24th International Conference on Computer Aided Verification", Berkeley, California, USA, M. PARTHASARATHY, S. A. SESHIA (editors), Lecture Notes in Computer Science, Springer, July 2012, vol. 7358, http://hal.archives-ouvertes.fr/hal-00799272

[6] S. CONCHON, M. IGUERNELALA, K. JI, G. MELQUIOND, C. FUMEX.*A Three-tier Strategy for Reasoning about Floating-Point Numbers in SMT*, in "Computer Aided Verification", 2017, https://hal.inria.fr/hal-01522770

[7] J.-C. FILLIÂTRE, L. GONDELMAN, A. PASKEVICH.*The Spirit of Ghost Code*, in "Formal Methods in System Design", 2016, vol. 48, n^{o} 3, p. 152–174, https://hal.archives-ouvertes.fr/hal-01396864v1

[8] C. FUMEX, C. DROSS, J. GERLACH, C. MARCHÉ.*Specification and Proof of High-Level Functional Properties of Bit-Level Programs*, in "8th NASA Formal Methods Symposium", Minneapolis, MN, USA, S. RAYADURGAM, O. TKACHUK (editors), Lecture Notes in Computer Science, Springer, June 2016, vol. 9690, p. 291–306, https://hal.inria.fr/hal-01314876

[9] C. MARCHÉ.*Verification of the Functional Behavior of a Floating-Point Program: an Industrial Case Study*, in "Science of Computer Programming", March 2014, vol. 96, n^{o} 3, p. 279–296, http://hal.inria.fr/hal-00967124/en

[10] É. MARTIN-DOREL, G. MELQUIOND.*Proving Tight Bounds on Univariate Expressions with Elementary Functions in Coq*, in "Journal of Automated Reasoning", 2016, https://hal.inria.fr/hal-01086460

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] M. CLOCHARD.*Methods and tools for specification and proof of difficult properties of sequential programs*, Université Paris-Saclay, March 2018, https://tel.archives-ouvertes.fr/tel-01787689

[12] D. DECLERCK.*Verification via Model Checking of Parameterized Concurrent Programs on Weak Memory Models*, Université Paris-Saclay, September 2018, https://tel.archives-ouvertes.fr/tel-01900842

Articles in International Peer-Reviewed Journal

[13] M. CLOCHARD, L. GONDELMAN, M. PEREIRA.*The Matrix Reproved: Verification Pearl*, in "Journal of Automated Reasoning", 2018, vol. 60, n° 3, p. 365-383 [*DOI :* 10.1007/S10817-017-9436-2], https://hal.inria.fr/hal-01617437

[14] S. DAILLER, D. HAUZAR, C. MARCHÉ, Y. MOY.*Instrumenting a Weakest Precondition Calculus for Counterexample Generation*, in "Journal of Logical and Algebraic Methods in Programming", 2018, vol. 99, p. 97–113, https://hal.inria.fr/hal-01802488

[15] A. MAHBOUBI, G. MELQUIOND, T. SIBUT-PINOTE.*Formally Verified Approximations of Definite Integrals*, in "Journal of Automated Reasoning", March 2018, p. 1-20 [*DOI :* 10.1007/S10817-018-9463-7], https://hal.inria.fr/hal-01630143

[16] A. VOLKOVA, M. ISTOAN, F. DE DINECHIN, T. HILAIRE.*Towards Hardware IIR Filters Computing Just Right: Direct Form I Case Study*, in "IEEE Transactions on Computers", 2018 [*DOI :* 10.1109/TC.2018.2879432], https://hal.sorbonne-universite.fr/hal-01561052

International Conferences with Proceedings

[17] S. BOLDO, F. FAISSOLE, V. TOURNEUR.*A Formally-Proved Algorithm to Compute the Correct Average of Decimal Floating-Point Numbers*, in "25th IEEE Symposium on Computer Arithmetic", Amherst, MA, United States, June 2018, https://hal.inria.fr/hal-01772272

[18] S. DAILLER, C. MARCHÉ, Y. MOY.*Lightweight Interactive Proving inside an Automatic Program Verifier*, in "4th Workshop on Formal Integrated Development Environment", Oxford, United Kingdom, P. MASCI, R. MONAHAN, V. PREVOSTO (editors), Electronic Proceedings in Theoretical Computer Science, Open Publishing Association, 2018, vol. 284, https://hal.inria.fr/hal-01936302

[19] D. GALLOIS-WONG, S. BOLDO, T. HILAIRE.*A Coq formalization of digital filters*, in "CICM 2018 - 11th Conference on Intelligent Computer Mathematics", Hagenberg, Austria, F. RABE, W. M. FARMER, G. O. PASSMORE, A. YOUSSEF (editors), Intelligent Computer Mathematics, August 2018, p. 87–103 [*DOI :* 10.1007/978-3-319-96812-4_8], https://hal.inria.fr/hal-01728828

[20] G. MELQUIOND, R. RIEU-HELFT.*A Why3 Framework for Reflection Proofs and its Application to GMP's Algorithms*, in "9th International Joint Conference on Automated Reasoning", Oxford, United Kingdom, D. GALMICHE, S. SCHULZ, R. SEBASTIANI (editors), Lecture Notes in Computer Science, July 2018, n° 10900, p. 178-193 [*DOI :* 10.1007/978-3-319-94205-6_13], https://hal.inria.fr/hal-01699754

[21] P. ROUX, M. IGUERNELALA, S. CONCHON.*A Non-linear Arithmetic Procedure for Control-Command Software Verification*, in "24th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS)", Thessalonique, Greece, Lecture Notes in Computer Science, Springer, April 2018, vol. 10806, p. 132-151, https://hal.archives-ouvertes.fr/hal-01737737

National Conferences with Proceeding

[22] F. FAISSOLE.*Définir le fini : deux formalisations d'espaces de dimension finie*, in "JLFA 2018 - Journées Francophones des Langages Applicatifs", Banyuls-sur-mer, France, 29èmes Journées Francophones des Langages Applicatifs, January 2018, p. 1-6, https://hal.inria.fr/hal-01654457

[23] F. FAISSOLE, B. SPITTERS.*Preuves constructives de programmes probabilistes*, in "JFLA 2018 - Journées Francophones des Langages Applicatifs", Banyuls-sur-Mer, France, 29èmes Journées Francophones des Langages Applicatifs, January 2018, https://hal.inria.fr/hal-01654459

[24] J.-C. FILLIÂTRE, M. PEREIRA, S. MELO DE SOUSA.*Vérification de programmes OCaml fortement impératifs avec Why3*, in "JFLA 2018 - Journées Francophones des Langages Applicatifs", Banyuls-sur-Mer, France, January 2018, p. 1-14, https://hal.inria.fr/hal-01649989

[25] D. GALLOIS-WONG.*Formalisation en Coq d'algorithmes de filtres numériques*, in "Journées Francophones des Langages Applicatifs (JFLA) 2019", Les Rousses, France, Journées Francophones des Langages Applicatifs 2019, Nicolas Magaud, January 2019, https://hal.inria.fr/hal-01929531

[26] R. RIEU-HELFT.*Un mécanisme d'extraction vers C pour Why3*, in "29èmes Journées Francophones des Langages Applicatifs", Banyuls-sur-Mer, France, January 2018, https://hal.inria.fr/hal-01653153

[27] R. RIEU-HELFT.*Un mécanisme de preuve par réflexion pour Why3 et son application aux algorithmes de GMP*, in "30èmes Journées Francophones des Langages Applicatifs", Rousses, France, January 2019, https://hal.inria.fr/hal-01943010

Scientific Books (or Scientific Book chapters)

[28] J.-M. MULLER, N. BRUNIE, F. DE DINECHIN, C.-P. JEANNEROD, M. JOLDES, V. LEFÈVRE, G. MELQUIOND, N. REVOL, S. TORRES.*Handbook of Floating-point Arithmetic (2nd edition)*, Birkhäuser Basel, July 2018, p. 1-627 [*DOI :* 10.1007/978-3-319-76526-6], https://hal.inria.fr/hal-01766584

Research Reports

[29] M. CLOCHARD, A. PASKEVICH, C. MARCHÉ.*Deductive Verification via Ghost Debugging*, Inria Saclay Ile de France, October 2018, n^{o} RR-9219, https://hal.inria.fr/hal-01907894

[30] Q. GARCHERY.*Démonstration automatique en Coq*, Paris Diderot ; Laboratoire de recherche en informatique (LRI) UMR CNRS 8623, Université Paris-Sud, August 2018, https://hal.archives-ouvertes.fr/hal-01874777

Other Publications

[31] S. BOLDO, F. FAISSOLE, A. CHAPOUTOT.*Round-off error and exceptional behavior analysis of explicit Runge-Kutta methods*, September 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01883843

[32] M. CLOCHARD, C. MARCHÉ, A. PASKEVICH.*Deductive Verification with Ghost Monitors*, November 2018, working paper or preprint, https://hal.inria.fr/hal-01926659

[33] J.-C. FILLIÂTRE, L. GONDELMAN, A. PASKEVICH, M. PEREIRA, S. MELO DE SOUSA.*A Toolchain to Produce Correct-by-Construction OCaml Programs*, May 2018, working paper or preprint, https://hal.inria.fr/hal-01783851

[34] D. GALLOIS-WONG, S. BOLDO, P. CUOQ.*Optimal Inverse Projection of Floating-Point Addition*, November 2018, working paper or preprint, https://hal.inria.fr/hal-01939097

[35] G.-A. JALOYAN, C. DROSS, M. MAALEJ, Y. MOY, A. PASKEVICH.*Verification of Programs with Pointers in SPARK*, November 2018, working paper or preprint, https://hal.inria.fr/hal-01936105

[36] A. VOLKOVA, T. HILAIRE, C. LAUTER.*Arithmetic approaches for rigorous design of reliable Fixed-Point LTI filters*, November 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01918650

References in notes

[37] B. BECKERT, R. HÄHNLE, P. H. SCHMITT (editors). *Verification of Object-Oriented Software: The KeY Approach*, Lecture Notes in Computer Science, Springer, 2007, vol. 4334

[38] U. A. ACAR, A. CHARGUÉRAUD, M. RAINEY.*Theory and Practice of Chunked Sequences*, in "European Symposium on Algorithms", Wroclaw, Poland, A. SCHULZ, D. WAGNER (editors), Lecture Notes in Computer Science, Springer, September 2014, vol. 8737, p. 25–36, https://hal.inria.fr/hal-01087245

[39] U. A. ACAR, A. CHARGUÉRAUD, M. RAINEY.*Oracle-Guided Scheduling for Controlling Granularity in Implicitly Parallel Languages*, in "Journal of Functional Programming", November 2016, vol. 26, https://hal.inria.fr/hal-01409069

[40] U. A. ACAR, A. CHARGUÉRAUD, M. RAINEY, F. SIECZKOWSKI.*Dag-calculus: a calculus for parallel computation*, in "Proceedings of the 21st ACM SIGPLAN International Conference on Functional Programming (ICFP)", Nara, Japan, September 2016, p. 18–32, https://hal.inria.fr/hal-01409022

[41] J. B. ALMEIDA, M. BARBOSA, J.-C. FILLIÂTRE, J. S. PINTO, B. VIEIRA.*CAOVerif: An Open-Source Deductive Verification Platform for Cryptographic Software Implementations*, in "Science of Computer Programming", October 2012

[42] A. AYAD, C. MARCHÉ.*Multi-Prover Verification of Floating-Point Programs*, in "Fifth International Joint Conference on Automated Reasoning", Edinburgh, Scotland, J. GIESL, R. HÄHNLE (editors), Lecture Notes in Artificial Intelligence, Springer, July 2010, vol. 6173, p. 127–141, http://hal.inria.fr/inria-00534333

[43] D. BAELDE, P. COURTIEU, D. GROSS-AMBLARD, C. PAULIN-MOHRING.*Towards Provably Robust Watermarking*, in "ITP 2012", Lecture Notes in Computer Science, August 2012, vol. 7406, http://hal.inria.fr/hal-00682398

[44] C. BARRETT, C. TINELLI.*CVC3*, in "19th International Conference on Computer Aided Verification", Berlin, Germany, W. DAMM, H. HERMANNS (editors), Lecture Notes in Computer Science, Springer, July 2007, vol. 4590, p. 298–302

[45] P. BAUDIN, J.-C. FILLIÂTRE, C. MARCHÉ, B. MONATE, Y. MOY, V. PREVOSTO.*ACSL: ANSI/ISO C Specification Language, version 1.4*, 2009

[46] P. BEHM, P. BENOIT, A. FAIVRE, J.-M. MEYNADIER.*METEOR : A successful application of B in a large project*, in "Proceedings of FM'99: World Congress on Formal Methods", J. M. WING, J. WOODCOCK, J. DAVIES (editors), Lecture Notes in Computer Science (Springer-Verlag), Springer Verlag, September 1999, p. 369–387

[47] J. C. BLANCHETTE, A. PASKEVICH.*TFF1: The TPTP typed first-order form with rank-1 polymorphism*, in "24th International Conference on Automated Deduction (CADE-24)", Lake Placid, USA, Lecture Notes in Artificial Intelligence, Springer, June 2013, vol. 7898, http://hal.inria.fr/hal-00825086

[48] F. BOBOT, S. CONCHON, É. CONTEJEAN, M. IGUERNELALA, S. LESCUYER, A. MEBSOUT.*The Alt-Ergo Automated Theorem Prover*, 2008

[49] F. BOBOT, J.-C. FILLIÂTRE.*Separation Predicates: a Taste of Separation Logic in First-Order Logic*, in "14th International Conference on Formal Ingineering Methods (ICFEM)", Kyoto, Japan, Lecture Notes in Computer Science, Springer, November 2012, vol. 7635, http://hal.inria.fr/hal-00825088

[50] F. BOBOT, J.-C. FILLIÂTRE, C. MARCHÉ, G. MELQUIOND, A. PASKEVICH.*Preserving User Proofs Across Specification Changes*, in "Verified Software: Theories, Tools, Experiments (5th International Conference VSTTE)", Atherton, USA, E. COHEN, A. RYBALCHENKO (editors), Lecture Notes in Computer Science, Springer, May 2013, vol. 8164, p. 191–201, http://hal.inria.fr/hal-00875395

[51] F. BOBOT, J.-C. FILLIÂTRE, C. MARCHÉ, G. MELQUIOND, A. PASKEVICH.*The Why3 platform, version 0.81*, version 0.81, LRI, CNRS & Univ. Paris-Sud & Inria Saclay, March 2013, http://hal.inria.fr/hal-00822856/

[52] M. BODIN, A. CHARGUÉRAUD, D. FILARETTI, P. GARDNER, S. MAFFEIS, D. NAUDZIUNIENE, A. SCHMITT, G. SMITH.*A Trusted Mechanised JavaScript Specification*, in "Proceedings of the 41st ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages", San Diego, USA, ACM Press, January 2014, http://hal.inria.fr/hal-00910135

[53] S. BOLDO.*How to Compute the Area of a Triangle: a Formal Revisit*, in "Proceedings of the 21th IEEE Symposium on Computer Arithmetic", Austin, Texas, USA, 2013, http://hal.inria.fr/hal-00790071

[54] S. BOLDO.*Deductive Formal Verification: How To Make Your Floating-Point Programs Behave*, Université Paris-Sud, October 2014, Thèse d'habilitation, https://hal.inria.fr/tel-01089643

[55] S. BOLDO.*Formal verification of tricky numerical computations*, in "16th GAMM-IMACS International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics", Würzburg, Germany, September 2014, https://hal.inria.fr/hal-01088692

[56] S. BOLDO, F. CLÉMENT, J.-C. FILLIÂTRE, M. MAYERO, G. MELQUIOND, P. WEIS.*Formal Proof of a Wave Equation Resolution Scheme: the Method Error*, in "Proceedings of the First Interactive Theorem Proving Conference", Edinburgh, Scotland, M. KAUFMANN, L. C. PAULSON (editors), LNCS, Springer, July 2010, vol. 6172, p. 147–162, http://hal.inria.fr/inria-00450789/

[57] S. BOLDO, F. CLÉMENT, J.-C. FILLIÂTRE, M. MAYERO, G. MELQUIOND, P. WEIS.*Trusting Computations: a Mechanized Proof from Partial Differential Equations to Actual Program*, in "Computers and Mathematics with Applications", 2014, vol. 68, n^{o} 3, p. 325–352, http://hal.inria.fr/hal-00769201

[58] S. BOLDO, F. FAISSOLE, A. CHAPOUTOT.*Round-off Error Analysis of Explicit One-Step Numerical Integration Methods*, in "24th IEEE Symposium on Computer Arithmetic", London, United Kingdom, July 2017, https://hal.archives-ouvertes.fr/hal-01581794

[59] S. BOLDO, J.-C. FILLIÂTRE, G. MELQUIOND.*Combining Coq and Gappa for Certifying Floating-Point Programs*, in "16th Symposium on the Integration of Symbolic Computation and Mechanised Reasoning", Grand Bend, Canada, Lecture Notes in Artificial Intelligence, Springer, July 2009, vol. 5625, p. 59–74

[60] S. BOLDO, J.-H. JOURDAN, X. LEROY, G. MELQUIOND.*A Formally-Verified C Compiler Supporting Floating-Point Arithmetic*, in "Proceedings of the 21th IEEE Symposium on Computer Arithmetic", Austin, Texas, USA, 2013, http://hal.inria.fr/hal-00743090

[61] S. BOLDO, J.-H. JOURDAN, X. LEROY, G. MELQUIOND.*Verified Compilation of Floating-Point Computations*, in "Journal of Automated Reasoning", February 2015, vol. 54, n^{o} 2, p. 135-163, https://hal.inria.fr/hal-00862689

[62] S. BOLDO, C. LELAY, G. MELQUIOND.*Improving Real Analysis in Coq: a User-Friendly Approach to Integrals and Derivatives*, in "Proceedings of the Second International Conference on Certified Programs and Proofs", Kyoto, Japan, C. HAWBLITZEL, D. MILLER (editors), Lecture Notes in Computer Science, December 2012, vol. 7679, p. 289–304, http://hal.inria.fr/hal-00712938

[63] S. BOLDO, C. LELAY, G. MELQUIOND.*Coquelicot: A User-Friendly Library of Real Analysis for Coq*, in "Mathematics in Computer Science", June 2015, vol. 9, n^{o} 1, p. 41-62, http://hal.inria.fr/hal-00860648

[64] S. BOLDO, C. LELAY, G. MELQUIOND.*Formalization of Real Analysis: A Survey of Proof Assistants and Libraries*, in "Mathematical Structures in Computer Science", 2016, http://hal.inria.fr/hal-00806920

[65] S. BOLDO, C. MARCHÉ.*Formal verification of numerical programs: from C annotated programs to mechanical proofs*, in "Mathematics in Computer Science", 2011, vol. 5, p. 377–393, http://hal.inria.fr/hal-00777605

[66] S. BOLDO, T. M. T. NGUYEN.*Proofs of numerical programs when the compiler optimizes*, in "Innovations in Systems and Software Engineering", 2011, vol. 7, p. 151–160, http://hal.inria.fr/hal-00777639

[67] T. BORMER, M. BROCKSCHMIDT, D. DISTEFANO, G. ERNST, J.-C. FILLIÂTRE, R. GRIGORE, M. HUISMAN, V. KLEBANOV, C. MARCHÉ, R. MONAHAN, W. MOSTOWSKI, N. POLIKARPOVA, C. SCHEBEN, G. SCHELLHORN, B. TOFAN, J. TSCHANNEN, M. ULBRICH.*The COST IC0701 Verification Competition 2011*, in "Formal Verification of Object-Oriented Software, Revised Selected Papers Presented at the International Conference, FoVeOOS 2011", B. BECKERT, F. DAMIANI, D. GUROV (editors), Lecture Notes in Computer Science, Springer, 2012, vol. 7421, http://hal.inria.fr/hal-00789525

[68] L. BURDY, Y. CHEON, D. R. COK, M. D. ERNST, J. R. KINIRY, G. T. LEAVENS, K. R. M. LEINO, E. POLL.*An overview of JML tools and applications*, in "International Journal on Software Tools for Technology Transfer (STTT)", June 2005, vol. 7, n^{o} 3, p. 212–232

[69] R. CHAPMAN, F. SCHANDA.*Are We There Yet? 20 Years of Industrial Theorem Proving with SPARK*, in "Interactive Theorem Proving - 5th International Conference, ITP 2014, Held as Part of the Vienna Summer of Logic, VSL 2014, Vienna, Austria, July 14-17, 2014. Proceedings", G. KLEIN, R. GAMBOA (editors), Lecture Notes in Computer Science, Springer, 2014, vol. 8558, p. 17–26

[70] A. CHARGUÉRAUD, F. POTTIER.*Verifying the Correctness and Amortized Complexity of a Union-Find Implementation in Separation Logic with Time Credits*, in "Journal of Automated Reasoning", September 2017

[71] M. CLOCHARD.*Automatically verified implementation of data structures based on AVL trees*, in "6th Working Conference on Verified Software: Theories, Tools and Experiments (VSTTE)", Vienna, Austria, D. GIANNAKOPOULOU, D. KROENING (editors), Lecture Notes in Computer Science, Springer, July 2014, vol. 8471, p. 167–180, http://hal.inria.fr/hal-01067217

[72] M. CLOCHARD, J.-C. FILLIÂTRE, C. MARCHÉ, A. PASKEVICH.*Formalizing Semantics with an Automatic Program Verifier*, in "6th Working Conference on Verified Software: Theories, Tools and Experiments (VSTTE)", Vienna, Austria, D. GIANNAKOPOULOU, D. KROENING (editors), Lecture Notes in Computer Science, Springer, July 2014, vol. 8471, p. 37–51, http://hal.inria.fr/hal-01067197

[73] M. CLOCHARD, L. GONDELMAN.*Double WP: vers une preuve automatique d'un compilateur*, in "Vingt-sixièmes Journées Francophones des Langages Applicatifs", Val d'Ajol, France, January 2015, https://hal.inria.fr/hal-01094488

[74] M. CLOCHARD, L. GONDELMAN, M. PEREIRA.*The Matrix Reproved*, in "8th Working Conference on Verified Software: Theories, Tools and Experiments (VSTTE)", Toronto, Canada, S. BLAZY, M. CHECHIK (editors), Lecture Notes in Computer Science, Springer, July 2016, https://hal.inria.fr/hal-01316902

[75] M. CLOCHARD, C. MARCHÉ, A. PASKEVICH.*Verified Programs with Binders*, in "Programming Languages meets Program Verification (PLPV)", ACM Press, 2014, http://hal.inria.fr/hal-00913431

[76] S. CONCHON.*SMT Techniques and their Applications: from Alt-Ergo to Cubicle*, Université Paris-Sud, December 2012, In English, http://www.lri.fr/~conchon/publis/conchonHDR.pdf, Thèse d'habilitation

[77] S. CONCHON, É. CONTEJEAN, M. IGUERNELALA.*Canonized Rewriting and Ground AC Completion Modulo Shostak Theories*, in "Tools and Algorithms for the Construction and Analysis of Systems", Saarbrücken, Germany, P. A. ABDULLA, K. R. M. LEINO (editors), Lecture Notes in Computer Science, Springer, April 2011, vol. 6605, p. 45-59, http://hal.inria.fr/hal-00777663

[78] S. CONCHON, É. CONTEJEAN, M. IGUERNELALA.*Canonized Rewriting and Ground AC Completion Modulo Shostak Theories : Design and Implementation*, in "Logical Methods in Computer Science", September 2012, vol. 8, n^o^ 3, p. 1–29, http://hal.inria.fr/hal-00798082

[79] S. CONCHON, D. DECLERCK, L. MARANGET, A. MEBSOUT.*Vérification de programmes C concurrents avec Cubicle : Enfoncer les barrières*, in "Vingt-cinquièmes Journées Francophones des Langages Applicatifs", Fréjus, France, January 2014, https://hal.inria.fr/hal-01088655

[80] S. CONCHON, A. GOEL, S. KRSTIĆ, A. MEBSOUT, F. ZAÏDI.*Invariants for Finite Instances and Beyond*, in "FMCAD", Portland, Oregon, États-Unis, October 2013, p. 61–68, http://hal.archives-ouvertes.fr/hal-00924640

[81] S. CONCHON, M. IGUERNELALA.*Tuning the Alt-Ergo SMT Solver for B Proof Obligations*, in "Abstract State Machines, Alloy, B, VDM, and Z (ABZ)", Toulouse, France, Lecture Notes in Computer Science, Springer, June 2014, vol. 8477, p. 294–297, https://hal.inria.fr/hal-01093000

[82] S. CONCHON, M. IGUERNELALA, A. MEBSOUT.*A Collaborative Framework for Non-Linear Integer Arithmetic Reasoning in Alt-Ergo*, 2013, https://hal.archives-ouvertes.fr/hal-00924646

[83] S. CONCHON, A. MEBSOUT, F. ZAÏDI.*Vérification de systèmes paramétrés avec Cubicle*, in "Vingt-quatrièmes Journées Francophones des Langages Applicatifs", Aussois, France, February 2013, http://hal.inria.fr/hal-00778832

[84] S. CONCHON, G. MELQUIOND, C. ROUX, M. IGUERNELALA.*Built-in Treatment of an Axiomatic Floating-Point Theory for SMT Solvers*, in "SMT workshop", Manchester, UK, P. FONTAINE, A. GOEL (editors), LORIA, 2012, p. 12–21

[85] M. DAHLWEID, M. MOSKAL, T. SANTEN, S. TOBIES, W. SCHULTE.*VCC: Contract-based modular verification of concurrent C*, in "31st International Conference on Software Engineering, ICSE 2009, May 16-24, 2009, Vancouver, Canada, Companion Volume", IEEE Comp. Soc. Press, 2009, p. 429-430

[86] D. DELAHAYE, C. DUBOIS, C. MARCHÉ, D. MENTRÉ.*The BWare Project: Building a Proof Platform for the Automated Verification of B Proof Obligations*, in "Abstract State Machines, Alloy, B, VDM, and Z (ABZ)", Toulouse, France, Lecture Notes in Computer Science, Springer, June 2014, vol. 8477, p. 290–293, http://hal.inria.fr/hal-00998092/en/

[87] D. DELAHAYE, C. MARCHÉ, D. MENTRÉ.*Le projet BWare : une plate-forme pour la vérification automatique d'obligations de preuve B*, in "Approches Formelles dans l'Assistance au Développement de Logiciels (AFADL)", Paris, France, EasyChair, June 2014, http://hal.inria.fr/hal-00998094/en/

[88] C. DROSS, S. CONCHON, J. KANIG, A. PASKEVICH.*Reasoning with Triggers*, in "SMT workshop", Manchester, UK, P. FONTAINE, A. GOEL (editors), LORIA, 2012

[89] C. DROSS.*Generic Decision Procedures for Axiomatic First-Order Theories*, Université Paris-Sud, April 2014, http://tel.archives-ouvertes.fr/tel-01002190

[90] F. FAISSOLE.*Formalization and closedness of finite dimensional subspaces*, in "SYNASC 2017 - 19th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing", Timioara, Romania, September 2017

[91] F. FAISSOLE, B. SPITTERS.*Synthetic topology in homotopy type theory for probabilistic programming*, January 2017, PPS 2017 - Workshop on probabilistic programming semantics, https://hal.inria.fr/hal-01485397

[92] F. FAISSOLE, B. SPITTERS.*Synthetic topology in HoTT for probabilistic programming*, in "The Third International Workshop on Coq for Programming Languages (CoqPL 2017)", Paris, France, January 2017, https://hal.inria.fr/hal-01405762

[93] J.-C. FILLIÂTRE.*Combining Interactive and Automated Theorem Proving in Why3 (invited talk)*, in "Automation in Proof Assistants 2012", Tallinn, Estonia, K. HELJANKO, H. HERBELIN (editors), April 2012

[94] J.-C. FILLIÂTRE.*Combining Interactive and Automated Theorem Proving using Why3 (invited tutorial)*, in "Second International Workshop on Intermediate Verification Languages (BOOGIE 2012)", Berkeley, California, USA, Z. RAKAMARIĆ (editor), July 2012

[95] J.-C. FILLIÂTRE. *Verifying Two Lines of C with Why3: an Exercise in Program Verification*, in "Verified Software: Theories, Tools, Experiments (4th International Conference VSTTE)", Philadelphia, USA, R. JOSHI, P. MÜLLER, A. PODELSKI (editors), Lecture Notes in Computer Science, Springer, January 2012, vol. 7152, p. 83–97

[96] J.-C. FILLIÂTRE. *Deductive Program Verification*, in "Programming Languages Mentoring Workshop (PLMW 2013)", Rome, Italy, N. FOSTER, P. GARDNER, A. SCHMITT, G. SMITH, P. THIEMAN, T. WRIGSTAD (editors), January 2013, http://hal.inria.fr/hal-00799190

[97] J.-C. FILLIÂTRE. *One Logic To Use Them All*, in "24th International Conference on Automated Deduction (CADE-24)", Lake Placid, USA, Lecture Notes in Artificial Intelligence, Springer, June 2013, vol. 7898, p. 1–20, http://hal.inria.fr/hal-00809651/en/

[98] J.-C. FILLIÂTRE, A. PASKEVICH. *Why3 — Where Programs Meet Provers*, in "Proceedings of the 22nd European Symposium on Programming", M. FELLEISEN, P. GARDNER (editors), Lecture Notes in Computer Science, Springer, March 2013, vol. 7792, p. 125–128, http://hal.inria.fr/hal-00789533

[99] J.-C. FILLIÂTRE, A. PASKEVICH, A. STUMP. *The 2nd Verified Software Competition: Experience Report*, in "COMPARE2012: 1st International Workshop on Comparative Empirical Evaluation of Reasoning Systems", Manchester, UK, V. KLEBANOV, S. GREBING (editors), EasyChair, June 2012, http://hal.inria.fr/hal-00798777

[100] J.-C. FILLIÂTRE, M. PEREIRA. *Producing All Ideals of a Forest, Formally (Verification Pearl)*, in "8th Working Conference on Verified Software: Theories, Tools and Experiments (VSTTE)", Toronto, Canada, S. BLAZY, M. CHECHIK (editors), Lecture Notes in Computer Science, Springer, July 2016, https://hal.inria.fr/hal-01316859

[101] D. HAUZAR, C. MARCHÉ, Y. MOY. *Counterexamples from Proof Failures in SPARK*, in "Software Engineering and Formal Methods", Vienna, Austria, R. DE NICOLA, E. KÜHN (editors), Lecture Notes in Computer Science, 2016, p. 215–233, https://hal.inria.fr/hal-01314885

[102] D. HAUZAR, C. MARCHÉ, Y. MOY. *Counterexamples from proof failures in the SPARK program verifier*, Inria Saclay Ile-de-France, February 2016, n^{o} RR-8854, https://hal.inria.fr/hal-01271174

[103] P. HERMS. *Certification of a Tool Chain for Deductive Program Verification*, Université Paris-Sud, January 2013, http://tel.archives-ouvertes.fr/tel-00789543

[104] P. HERMS, C. MARCHÉ, B. MONATE. *A Certified Multi-prover Verification Condition Generator*, in "Verified Software: Theories, Tools, Experiments (4th International Conference VSTTE)", Philadelphia, USA, R. JOSHI, P. MÜLLER, A. PODELSKI (editors), Lecture Notes in Computer Science, Springer, January 2012, vol. 7152, p. 2–17, http://hal.inria.fr/hal-00639977

[105] M. IGUERNELALA. *Strengthening the Heart of an SMT-Solver: Design and Implementation of Efficient Decision Procedures*, Université Paris-Sud, June 2013, http://tel.archives-ouvertes.fr/tel-00842555

[106] D. ISHII, G. MELQUIOND, S. NAKAJIMA. *Inductive Verification of Hybrid Automata with Strongest Postcondition Calculus*, in "Proceedings of the 10th Conference on Integrated Formal Methods", Turku, Finland, E. B. JOHNSEN, L. PETRE (editors), Lecture Notes in Computer Science, 2013, vol. 7940, p. 139–153, http://hal.inria.fr/hal-00806701

[107] J. KANIG, E. SCHONBERG, C. DROSS.*Hi-Lite: the convergence of compiler technology and program verification*, in "Proceedings of the 2012 ACM Conference on High Integrity Language Technology, HILT '12", Boston, USA, B. BROSGOL, J. BOLENG, S. T. TAFT (editors), ACM Press, 2012, p. 27–34

[108] G. KLEIN, J. ANDRONICK, K. ELPHINSTONE, G. HEISER, D. COCK, P. DERRIN, D. ELKADUWE, K. ENGELHARDT, R. KOLANSKI, M. NORRISH, T. SEWELL, H. TUCH, S. WINWOOD.*seL4: Formal verification of an OS kernel*, in "Communications of the ACM", June 2010, vol. 53, n° 6, p. 107–115

[109] C. LELAY.*A New Formalization of Power Series in Coq*, in "5th Coq Workshop", Rennes, France, July 2013, p. 1–2, http://hal.inria.fr/hal-00880212

[110] C. LELAY.*Coq passe le bac*, in "JFLA - Journées francophones des langages applicatifs", Fréjus, France, January 2014

[111] C. LELAY, G. MELQUIOND.*Différentiabilité et intégrabilité en Coq. Application à la formule de d'Alembert*, in "Vingt-troisièmes Journées Francophones des Langages Applicatifs", Carnac, France, February 2012, http://hal.inria.fr/hal-00642206/fr/

[112] X. LEROY.*A formally verified compiler back-end*, in "Journal of Automated Reasoning", 2009, vol. 43, n° 4, p. 363–446, http://hal.inria.fr/inria-00360768/en/

[113] A. MAHBOUBI, G. MELQUIOND, T. SIBUT-PINOTE.*Formally Verified Approximations of Definite Integrals*, in "Proceedings of the 7th International Conference on Interactive Theorem Proving", Nancy, France, J. C. BLANCHETTE, S. MERZ (editors), Lecture Notes in Computer Science, August 2016, vol. 9807, https://hal.inria.fr/hal-01289616

[114] C. MARCHÉ, A. TAFAT.*Weakest Precondition Calculus, revisited using Why3*, Inria, December 2012, n° RR-8185, http://hal.inria.fr/hal-00766171

[115] C. MARCHÉ, A. TAFAT.*Calcul de plus faible précondition, revisité en Why3*, in "Vingt-quatrièmes Journées Francophones des Langages Applicatifs", Aussois, France, February 2013, http://hal.inria.fr/hal-00778791

[116] C. MARCHÉ.*Verification of the Functional Behavior of a Floating-Point Program: an Industrial Case Study*, in "Science of Computer Programming", March 2014, vol. 96, n° 3, p. 279–296, http://hal.inria.fr/hal-00967124/en

[117] É. MARTIN-DOREL, G. MELQUIOND, J.-M. MULLER.*Some Issues related to Double Roundings*, in "BIT Numerical Mathematics", 2013, vol. 53, n° 4, p. 897–924, http://hal-ens-lyon.archives-ouvertes.fr/ensl-00644408

[118] A. MEBSOUT.*Invariants inference for model checking of parameterized systems*, Université Paris-Sud, September 2014, https://tel.archives-ouvertes.fr/tel-01073980

[119] G. MELQUIOND.*Floating-point arithmetic in the Coq system*, in "Information and Computation", 2012, vol. 216, p. 14–23, http://hal.inria.fr/hal-00797913

[120] G. MELQUIOND, W. G. NOWAK, P. ZIMMERMANN.*Numerical Approximation of the Masser-Gramain Constant to Four Decimal Digits: delta=1.819...*, in "Mathematics of Computation", 2013, vol. 82, p. 1235–1246, http://hal.inria.fr/hal-00644166/en/

[121] D. MENTRÉ, C. MARCHÉ, J.-C. FILLIÂTRE, M. ASUKA.*Discharging Proof Obligations from Atelier B using Multiple Automated Provers*, in "ABZ'2012 - 3rd International Conference on Abstract State Machines, Alloy, B and Z", Pisa, Italy, S. REEVES, E. RICCOBENE (editors), Lecture Notes in Computer Science, Springer, June 2012, vol. 7316, p. 238–251, http://hal.inria.fr/hal-00681781/en/

[122] J.-M. MULLER, N. BRISEBARRE, F. DE DINECHIN, C.-P. JEANNEROD, V. LEFÈVRE, G. MELQUIOND, N. REVOL, D. STEHLÉ, S. TORRES.*Handbook of Floating-Point Arithmetic*, Birkhäuser, 2010

[123] T. M. T. NGUYEN, C. MARCHÉ.*Hardware-Dependent Proofs of Numerical Programs*, in "Certified Programs and Proofs", J.-P. JOUANNAUD, Z. SHAO (editors), Lecture Notes in Computer Science, Springer, December 2011, p. 314–329, http://hal.inria.fr/hal-00772508

[124] T. M. T. NGUYEN.*Taking architecture and compiler into account in formal proofs of numerical programs*, Université Paris-Sud, June 2012, http://tel.archives-ouvertes.fr/tel-00710193

[125] M. NORRISH.*C Formalised in HOL*, University of Cambridge, November 1998

[126] M. PEREIRA, J.-C. FILLIÂTRE, S. M. DE SOUSA.*ARMY: a Deductive Verification Platform for ARM Programs Using Why3*, in "INForum 2012", September 2012

[127] M. PEREIRA.*Tools and Techniques for the Verification of Modular Stateful Code*, Université Paris-Saclay, December 2018

[128] R. RIEU-HELFT, C. MARCHÉ, G. MELQUIOND.*How to Get an Efficient yet Verified Arbitrary-Precision Integer Library*, in "9th Working Conference on Verified Software: Theories, Tools, and Experiments", Heidelberg, Germany, Lecture Notes in Computer Science, July 2017, vol. 10712, p. 84–101, https://hal.inria.fr/hal-01519732

[129] P. ROUX.*Formal Proofs of Rounding Error Bounds*, in "Journal of Automated Reasoning", 2015, https://hal.archives-ouvertes.fr/hal-01091189

[130] N. SCHIRMER.*Verification of Sequential Imperative Programs in Isabelle/HOL*, Technische Universität München, 2006

[131] A. TAFAT.*Preuves par raffinement de programmes avec pointeurs*, Université Paris-Sud, September 2013, http://tel.archives-ouvertes.fr/tel-00874679

[132] F. DE DINECHIN, C. LAUTER, G. MELQUIOND.*Certifying the floating-point implementation of an elementary function using Gappa*, in "IEEE Transactions on Computers", 2011, vol. 60, n° 2, p. 242–253, http://hal.inria.fr/inria-00533968/en/

[133] L. DE MOURA, N. BJØRNER.*Z3, An Efficient SMT Solver*, in "TACAS", Lecture Notes in Computer Science, Springer, 2008, vol. 4963, p. 337–340

Project-Team TROPICAL

Tropical methods: structures, algorithms and interactions

IN COLLABORATION WITH: Centre de Mathématiques Appliquées (CMAP)

RESEARCH CENTER
Saclay - Île-de-France

THEME
Optimization and control of dynamic systems

Table of contents

Project-Team TROPICAL

Creation of the Team: 2016 January 01, updated into Project-Team: 2018 July 01

Keywords:

Computer Science and Digital Science:

A1.2.4. - QoS, performance evaluation
A2.3.3. - Real-time systems
A2.4. - Formal method for verification, reliability, certification
A6.2.5. - Numerical Linear Algebra
A6.2.6. - Optimization
A6.4.6. - Optimal control
A7.2.4. - Mechanized Formalization of Mathematics
A8.1. - Discrete mathematics, combinatorics
A8.2.1. - Operations research
A8.2.3. - Calculus of variations
A8.3. - Geometry, Topology
A8.9. - Performance evaluation
A8.11. - Game Theory
A9.6. - Decision support

Other Research Topics and Application Domains:

B4.3. - Renewable energy production
B4.4. - Energy delivery
B4.4.1. - Smart grids
B6.6. - Embedded systems
B8.4. - Security and personal assistance
B8.4.1. - Crisis management

1. Team, Visitors, External Collaborators

Research Scientists

Stéphane Gaubert [Team leader, Inria, Senior Researcher, HDR]
Marianne Akian [Inria, Senior Researcher, HDR]
Xavier Allamigeon [Corps des Mines, under secondment, Inria, Researcher]
Cormac Walsh [Inria, Researcher]

Technical Staff

Nikolas Stott [Inria, until Feb 2018]

PhD Students

Marin Boyet [Inria, from Oct 2018]
Jean-Bernard Eytard [Inria]
Eric Fodjo [I-Fihn Consulting, until Jul 2018]
Maxime Grangereau [EDF Labs (Cifre)]
Paulin Jacquot [EDF Labs (Cifre)]
Omar Saadi [Ecole polytechnique, from Sep 2018]
Mateusz Skomra [Ecole polytechnique, until Sep 2018]

Duy Nghi Benoît Tran [Univ Paris-Est Marne La Vallée]

Post-Doctoral Fellows

Daniel Jones [Inria, Post-Doctoral Fellow, until Mar 2018]

Aurélien Sagnier [Ecole polytechnique]

Visiting Scientists

Roberto Bobadilla Solari [DIM and CMM, University of Chili, Santiago, from Sep 2018 until Nov 2018]

Aris Daniilidis [DIM and CMM, University of Chili, Santiago, Sep 2018, HDR]

Gonzalo Flores Garcia [DIM and CMM, University of Chili, Santiago, from Sep 2018]

Ricardo Katz [CONICET, from May 2018 until Jul 2018]

Gleb Koshevoy [Russian Academy of Sciences, Nov 2018]

Sebastian Tapia Garcia [DIM and CMM, University of Chili, Santiago, from Sep 2018]

Administrative Assistants

Hanadi Dib [Inria]

Ines Dumontier [Inria, from Nov 2018]

2. Overall Objectives

2.1. Introduction

The project develops tropical methods motivated by applications arising in decision theory (deterministic and stochastic optimal control, game theory, optimization and operations research), in the analysis or control of classes of dynamical systems (including timed discrete event systems and positive systems), in the verification of programs and systems, and in the development of numerical algorithms. Tropical algebra tools are used in interaction with various methods, coming from convex analysis, Hamilton–Jacobi partial differential equations, metric geometry, Perron-Frobenius and nonlinear fixed-point theories, combinatorics or algorithmic complexity. The emphasis of the project is on mathematical modelling and computational aspects.

The subtitle of the *Tropical* project, namely, “structures, algorithms, and interactions”, refers to the spirit of our research, including a methodological component, computational aspects, and finally interactions with other scientific fields or real world applications, in particular through mathematical modelling.

2.2. Scientific context

Tropical algebra, geometry, and analysis have enjoyed spectacular development in recent years. Tropical structures initially arose to solve problems in performance evaluation of discrete event systems [62], combinatorial optimization [67], or automata theory [117]. They also arose in mathematical physics and asymptotic analysis [107], [104]. More recently, these structures have appeared in several areas of pure mathematics, in particular in the study of combinatorial aspects of algebraic geometry [96], [131], [120], [101], in algebraic combinatorics [88], and in arithmetics [73]. Also, further applications of tropical methods have appeared, including optimal control [111], program invariant computation [57] and timed systems verification [106], and zero-sum games [2].

The term ‘tropical’ generally refers to algebraic structures in which the laws originate from optimization processes. The prototypical tropical structure is the max-plus semifield, consisting of the real numbers, equipped with the maximum, thought of as an additive law, and the addition, thought of as a multiplicative law. Tropical objects appear as limits of classical objects along certain deformations (“log-limits sets” of Bergman, “Maslov dequantization”, or “Viro deformation”). For this reason, the introduction of tropical tools often yields new insights into old familiar problems, leading either to counterexamples or to new methods and results; see for instance [131], [113]. In some applications, like optimal control, discrete event systems, or static analysis of programs, tropical objects do not appear through a limit procedure, but more directly as a modelling or computation/analysis tool; see for instance [127], [62], [99], [68].

Tropical methods are linked to the fields of positive systems and of metric geometry [115], [11]. Indeed, tropically linear maps are monotone (a.k.a. order-preserving). They are also nonexpansive in certain natural metrics (sup-norm, Hopf oscillation, Hilbert's projective metric, ...). In this way, tropical dynamical systems appear to be special cases of nonexpansive, positive, or monotone dynamical systems, which are studied as part of linear and non-linear Perron-Frobenius theory [105], [3]. Such dynamical systems are of fundamental importance in the study of repeated games [112]. Monotonicity properties are also essential in the understanding of the fixed points problems which determine program invariants by abstract interpretation [75]. The latter problems are actually somehow similar to the ones arising in the study of zero-sum games; see [7]. Moreover, positivity or monotonicity methods are useful in population dynamics, either in a discrete space setting [129] or in a PDE setting [63]. In such cases, solving tropical problems often leads to solutions or combinatorial insights on classical problems involving positivity conditions (e.g., finding equilibria of dynamical systems with nonnegative coordinates, understanding the qualitative and quantitative behavior of growth rates / Floquet eigenvalues [9], etc). Other applications of Perron-Frobenius theory originate from quantum information and control [119], [125].

3. Research Program

3.1. Optimal control and zero-sum games

The dynamic programming approach allows one to analyze one or two-player dynamic decision problems by means of operators, or partial differential equations (Hamilton–Jacobi or Isaacs PDEs), describing the time evolution of the value function, i.e., of the optimal reward of one player, thought of as a function of the initial state and of the horizon. We work especially with problems having long or infinite horizon, modelled by stopping problems, or ergodic problems in which one optimizes a mean payoff per time unit. The determination of optimal strategies reduces to solving nonlinear fixed point equations, which are obtained either directly from discrete models, or after a discretization of a PDE.

The geometry of solutions of optimal control and game problems Basic questions include, especially for stationary or ergodic problems, the understanding of existence and uniqueness conditions for the solutions of dynamic programming equations, for instance in terms of controllability or ergodicity properties, and more generally the understanding of the structure of the full set of solutions of stationary Hamilton–Jacobi PDEs and of the set of optimal strategies. These issues are already challenging in the one-player deterministic case, which is an application of choice of tropical methods, since the Lax-Oleinik semigroup, i.e., the evolution semigroup of the Hamilton-Jacobi PDE, is a linear operator in the tropical sense. Recent progress in the deterministic case has been made by combining dynamical systems and PDE techniques (weak KAM theory [85]), and also using metric geometry ideas (abstract boundaries can be used to represent the sets of solutions [100], [4]). The two player case is challenging, owing to the lack of compactness of the analogue of the Lax-Oleinik semigroup and to a richer geometry. The conditions of solvability of ergodic problems for games (for instance, solvability of ergodic Isaacs PDEs), and the representation of solutions are only understood in special cases, for instance in the finite state space case, through tropical geometry and non-linear Perron-Frobenius methods [47],[14], [3].

Algorithmic aspects: from combinatorial algorithms to the attenuation of the curse of dimensionality Our general goal is to push the limits of solvable models by means of fast algorithms adapted to large scale instances. Such instances arise from discrete problems, in which the state space may so large that it is only accessible through local oracles (for instance, in some web ranking applications, the number of states may be the number of web pages) [86]. They also arise from the discretization of PDEs, in which the number of states grows exponentially with the number of degrees of freedom, according to the "curse of dimensionality". A first line of research is the development of *new approximation methods for the value function*. So far, classical approximations by linear combinations have been used, as well as approximation by suprema of linear or quadratic forms, which have been introduced in the setting of dual dynamic programming and of the so called "max-plus basis methods" [87]. We believe that more concise or more accurate approximations may be obtained by unifying these methods. Also, some max-plus basis methods have been shown to *attenuate the curse of dimensionality* for very special problems (for instance involving switching) [108], [92]. This suggests that the complexity of control or games problems may be measured by more subtle quantities that the mere number of states, for instance, by some forms of metric entropy (for example, certain large scale problems have a low complexity owing to the presence of decomposition properties, “highway hierarchies”, etc.). A second line of of our research is the development of *combinatorial algorithms*, to solve large scale zero-sum two-player problems with discrete state space. This is related to current open problems in algorithmic game theory. In particular, the existence of polynomial-time algorithms for games with ergodic payment is an open question. See e.g. [53] for a polynomial time average complexity result derived by tropical methods. The two lines of research are related, as the understanding of the geometry of solutions allows to develop better approximation or combinatorial algorithms.

3.2. Non-linear Perron-Frobenius theory, nonexpansive mappings and metric geometry

Several applications (including population dynamics [9] and discrete event systems [62], [70], [55]) lead to studying classes of dynamical systems with remarkable properties: preserving a cone, preserving an order, or being nonexpansive in a metric. These can be studied by techniques of non-linear Perron-Frobenius theory [3] or metric geometry [10]. Basic issues concern the existence and computation of the “escape rate” (which determines the throughput, the growth rate of the population), the characterizations of stationary regimes (non-linear fixed points), or the study of the dynamical properties (convergence to periodic orbits). Nonexpansive mappings also play a key role in the “operator approach” to zero-sum games, since the one-day operators of games are nonexpansive in several metrics, see [8].

3.3. Tropical algebra and convex geometry

The different applications mentioned in the other sections lead us to develop some basic research on tropical algebraic structures and in convex and discrete geometry, looking at objects or problems with a “piecewise-linear ” structure. These include the geometry and algorithmics of tropical convex sets [58], [50], tropical semialgebraic sets [60], the study of semi-modules (analogues of vector spaces when the base field is replaced by a semi-field), the study of systems of equations linear in the tropical sense, investigating for instance the analogues of the notions of rank, the analogue of the eigenproblems [52], and more generally of systems of tropical polynomial equations. Our research also builds on, and concern, classical convex and discrete geometry methods.

3.4. Tropical methods applied to optimization, perturbation theory and matrix analysis

Tropical algebraic objects appear as a deformation of classical objects thought various asymptotic procedures. A familiar example is the rule of asymptotic calculus,

$$e^{-a/\epsilon} + e^{-b/\epsilon} \asymp e^{-\min(a,b)/\epsilon} , \qquad e^{-a/\epsilon} \times e^{-b/\epsilon} = e^{-(a+b)/\epsilon} , \tag{16}$$

when $\epsilon \to 0^+$. Deformations of this kind have been studied in different contexts: large deviations, zero-temperature limits, Maslov’s “dequantization method” [107], non-archimedean valuations, log-limit sets and Viro’s patchworking method [132], etc.

This entails a relation between classical algorithmic problems and tropical algorithmic problems, one may first solve the $\epsilon = 0$ case (non-archimedean problem), which is sometimes easier, and then use the information gotten in this way to solve the $\epsilon = 1$ (archimedean) case.

In particular, tropicalization establishes a connection between polynomial systems and piecewise affine systems that are somehow similar to the ones arising in game problems. It allows one to transfer results from the world of combinatorics to “classical” equations solving. We investigate the consequences of this correspondence on complexity and numerical issues. For instance, combinatorial problems can be solved in a robust way. Hence, situations in which the tropicalization is faithful lead to improved algorithms for classical problems. In particular, scalings for the polynomial eigenproblems based on tropical preprocessings have started to be used in matrix analysis [93], [98].

Moreover, the tropical approach has been recently applied to construct examples of linear programs in which the central path has an unexpectedly high total curvature [54], and it has also led to positive polynomial-time average case results concerning the complexity of mean payoff games. Similarly, we are studying semidefinite programming over non-archimedean fields [60], [59], with the goal to better understand complexity issues in classical semidefinite and semi-algebraic programming.

4. Application Domains

4.1. Discrete event systems (manufacturing systems, networks)

One important class of applications of max-plus algebra comes from discrete event dynamical systems [62]. In particular, modelling timed systems subject to synchronization and concurrency phenomena leads to studying dynamical systems that are non-smooth, but which have remarkable structural properties (nonexpansiveness in certain metrics , monotonicity) or combinatorial properties. Algebraic methods allow one to obtain analytical expressions for performance measures (throughput, waiting time, etc). A recent application, to emergency call centers, can be found in [55].

4.2. Optimal control and games

Optimal control and game theory have numerous well established applications fields: mathematical economy and finance, stock optimization, optimization of networks, decision making, etc. In most of these applications, one needs either to derive analytical or qualitative properties of solutions, or design exact or approximation algorithms adapted to large scale problems.

4.3. Operations Research

We develop, or have developed, several aspects of operations research, including the application of stochastic control to optimal pricing, optimal measurement in networks [122]. Applications of tropical methods arise in particular from discrete optimization [68], [69], scheduling problems with and-or constraints [114], or product mix auctions [130].

4.4. Computing program and dynamical systems invariants

A number of programs and systems verification questions, in which safety considerations are involved, reduce to computing invariant subsets of dynamical systems. This approach appears in various guises in computer science, for instance in static analysis of program by abstract interpretation, along the lines of P. and R. Cousot [75], but also in control (eg, computing safety regions by solving Isaacs PDEs). These invariant sets are often sought in some tractable effective class: ellipsoids, polyhedra, parametric classes of polyhedra with a controlled complexity (the so called “templates” introduced by Sankaranarayanan, Sipma and Manna [123]), shadows of sets represented by linear matrix inequalities, disjunctive constraints represented by tropical polyhedra [57], etc. The computation of invariants boils down to solving large scale fixed point problems. The latter are of the same nature as the ones encountered in the theory of zero-sum games, and so, the techniques developed in the previous research directions (especially methods of monotonicity, nonexpansiveness, discretization of PDEs, etc) apply to the present setting, see e.g. [90], [95] for the application of policy iteration type algorithms, or for the application for fixed point problems over the space of quadratic forms [7]. The problem of computation of invariants is indeed a key issue needing the methods of several fields: convex and nonconvex programming, semidefinite programming and symbolic computation (to handle semialgebraic invariants), nonlinear fixed point theory, approximation theory, tropical methods (to handle disjunctions), and formal proof (to certify numerical invariants or inequalities).

5. Highlights of the Year

5.1. Highlights of the Year

- The paper [89] has been included in a list of "10 notable papers published over the last 50 years by the journal Linear Algebra and its applications", at the occasion of the golden anniversary of the journal.
- The article [17] answers an old question in the theory of interior point methods: it provides a counter example showing that log-barrier interior point methods are not strongly polynomial.

6. New Software and Platforms

6.1. Coq-Polyhedra

KEYWORDS: Coq - Polyhedra - Automated theorem proving - Linear optimization
SCIENTIFIC DESCRIPTION: Coq-Polyhedra is a library providing a formalization of convex polyhedra in the Coq proof assistant. While still in active development, it provides an implementation of the simplex method, and already handles the basic properties of polyhedra such as emptiness, boundedness, membership. Several fundamental results in the theory of convex polyhedra, such as Farkas Lemma, duality theorem of linear programming, and Minkowski Theorem, are also formally proved.

The formalization is based on the Mathematical Components library, and makes an extensive use of the boolean reflection methodology.
FUNCTIONAL DESCRIPTION: Coq-Polyhedra is a library which aims at formalizing convex polyhedra in Coq

- Participants: Xavier Allamigeon, Vasileios Charisopoulos and Ricardo Katz
- Partner: CIFASIS
- Contact: Xavier Allamigeon
- Publications: A Formalization of Convex Polyhedra Based on the Simplex Method - A Formalization of Convex Polyhedra Based on the Simplex Method - First steps in the formalization of convex polyhedra in Coq
- URL: https://github.com/nhojem/Coq-Polyhedra

7. New Results

7.1. Optimal control and zero-sum games

7.1.1. Fixed points of order preserving homogeneous maps and zero-sum games

Participants: Marianne Akian, Stéphane Gaubert.

In a series of joint works with Antoine Hochart, we apply methods of non-linear fixed point theory to zero-sum games.

A key issue is the solvability of the ergodic equation associated to a zero-sum game with finite state space, i.e., given a dynamic programming operator T associated to an undiscounted problem, one looks for a vector u, called the bias, and for a scalar λ, the ergodic constant, such that $T(u) = \lambda e + u$. The bias vector is of interest as it allows to determine optimal stationnary strategies.

In [14], we studied zero-sum games with perfect information and finite action spaces, and showed that the set of payments for which the bias vector is not unique (up to an additive constant) coincides with the union of lower dimensional cells of a polyhedral complex, in particular, the bias vector is unique, generically. We provided an application to perturbation schemes in policy iteration.

In [36], we apply game theory methods to the study of the nonlinear eigenproblem for homogeneous order preserving self maps of the interior of the cone. We show that the existence and uniqueness of an eigenvector is governed by combinatorial conditions, involving dominions (sets of states "controlled" by one of the two players). In this way, we characterize the situation in which the existence of an eigenvector holds independently of perturbations, and we solve an open problem raised in [91].

In [15], we provide a representation theorem for "payment free" Shapley operators, showing that these are characterized by monotonicity and homogeneity axioms [15]. This extends to the two-player case known representation theorems for risk measures.

7.1.2. Nonlinear fixed point methods to compute joint spectral raddi of nonnegative matrices

Participants: Stéphane Gaubert, Nikolas Stott.

In [29], we introduce a non-linear fixed point method to approximate the joint spectral radius of a finite set of nonnegative matrices. We show in particular that the joint spectral radius is the limit of the eigenvalues of a family of non-linear risk-sensitive type dynamic programming operators. We develop a projective version of Krasnoselskii-Mann iteration to solve these eigenproblems, and report experimental results on large scale instances (several matrices in dimensions of order 1000 within a minute). The situation in which the matrices are not nonnegative is amenable to a similar approach [94].

7.1.3. Probabilistic and max-plus approximation of Hamilton-Jacobi-Bellman equations

Participants: Marianne Akian, Eric Fodjo.

The PhD thesis of Eric Fodjo concerns stochastic control problems obtained in particular in the modelisation of portfolio selection with transaction costs. The dynamic programming method leads to a Hamilton-Jacobi-Bellman partial differential equation, on a space with a dimension at least equal to the number of risky assets. The curse of dimensionality does not allow one to solve numerically these equations for a large dimension (greater to 5). We propose to tackle these problems with numerical methods combining policy iterations, probabilistic discretisations, max-plus discretisations, in order to increase the possible dimension.

We consider fully nonlinear Hamilton-Jacobi-Bellman equations associated to diffusion control problems with finite horizon involving a finite set-valued (or switching) control and possibly a continuum-valued control. In [46], we constructed a lower complexity probabilistic numerical algorithm by combining the idempotent expansion properties obtained by McEneaney, Kaise and Han [103], [109] for solving such problems with a numerical probabilistic method such as the one proposed by Fahim, Touzi and Warin [82] for solving some fully nonlinear parabolic partial differential equations, when the volatility does not oscillate too much. In [32], [33], we improve the method of Fahim, Touzi and Warin by introducing probabilistic schemes which are monotone without any restrictive condition, allowing one to solve fully nonlinear parabolic partial differential equations with general volatilities. We study the convergence and obtain error estimates when the parameters and the value function are bounded. The more general quadratic growth case has been studied in the PhD manuscript [12].

7.1.4. Tropical-SDDP algorithms for stochastic control problems involving a switching control

Participants: Marianne Akian, Duy Nghi Benoît Tran.

The PhD thesis of Benoît Tran, supervised by Jean-Philippe Chancelier (ENPC) and Marianne Akian concerns the numerical solution of the dynamic programming equation of discrete time stochastic control problems.

Several methods have been proposed in the litterature to bypass the curse of dimensionality difficulty of such an equation, by assuming a certain structure of the problem. Examples are the max-plus based method of McEneaney [110], [111], the stochastic max-plus scheme proposed by Zheng Qu [118], the stochastic dual dynamic programming (SDDP) algorithm of Pereira and Pinto [116], the mixed integer dynamic approximation scheme of Philpott, Faisal and Bonnans [61], the probabilistic numerical method of Fahim, Touzi and Warin [82]. We propose to associate and compare these methods in order to solve more general structures.

In a first work [35], we build a common framework for both the SDDP and a discrete time and finite horizon version of Zheng Qu's algorithm for deterministic problems involving a finite set-valued (or switching) control and a continuum-valued control. We propose an algorithm that generates monotone approximations of the value function as a pointwise supremum, or infimum, of basic (affine or quadratic for example) functions which are randomly selected. We give sufficient conditions that ensure almost sure convergence of the approximations to the value function.

7.1.5. Parametrized complexity of optimal control and zero-sum game problems

Participants: Marianne Akian, Stéphane Gaubert, Omar Saadi.

As already said above, the dynamic programing approach to optimal control and zero-sum game problems suffers of the curse of dimensionality. The aim of the PhD thesis is to unify different techniques to bypass this difficulty, in order to obtain new algorithms and new complexity results.

As a first step, we worked to extend an algorithm proposed by Sidford et al. in [126]. There, they proposed a randomized value iteration algorithm which improves the usual complexity bounds of the value iteration for *discounted* Markov Decision Problems (discrete time stochastic control problems). In a joint work with Zheng Qu (Hong Kong University), we are extending this algorithm to the ergodic (mean payoff) case, exploiting techniques from non-linear spectral theory [48]; this extension covers as well the case of two players (zero-sum).

7.2. Non-linear Perron-Frobenius theory, nonexpansive mappings and metric geometry

7.2.1. Order isomorphisms and antimorphisms on cones

Participant: Cormac Walsh.

We have been studying non-linear operators on open cones, particularly ones that preserve or reverse the order structure associated to the cone. A bijective map that preserves the order in both directions is called an order isomorphism. Those that reverse the order in both directions are order antimorphisms. These are closely related to the isometries of the Hilbert and Thompson metrics on the cone.

Previously, we have shown [133] that if there exists an antimorphism on a finite-dimensional open cone that is homogeneous of degree -1, then the cone must be a symmetric cone, that is, have a transitive group of linear automorphisms and be self-dual. This result was improved in [44], where we showed that the homogeneity assumption is not actually necessary: every antimorphism on a cone is automatically homogeneous of degree -1.

The study of the order isomorphisms of a cone goes back to Alexandrov and Zeeman, who considered maps preserving the light cone that arises in special relativity. This work was extended to more general cones by Rothaus; Noll and Schäffer; and Artstein-Avidan and Slomka. It was shown, in the finite-dimensional case, that all isomorphisms are linear if the cone has no one-dimensional factors. There are also some results in infinite dimension—however these are unsatisfactory because of the strong assumptions that must be made in order to get the finite-dimensional techniques to work. For example, a typical assumption is that the cone is the convex hull of its extreme rays, which is overly restrictive in infinite dimension. Using different techniques more suited to infinite dimension, we have been developing a necessary and sufficient criterion on the geometry of a cone for all its isomorphisms to be linear.

7.2.2. Horofunction compactifications of symmetric spaces

Participant: Cormac Walsh.

This work is in collaboration with Thomas Haettel (Montpellier), Anna-Sofie Schilling (Heidelberg), Anna Wienhard (Heidelberg).

The symmetric spaces form a fascinating class of geometrical space. These are the spaces in which there is a point reflection through every point. An example is the space $\mathrm{Pos}(\mathbb{C}, n)$ of positive definite $n \times n$ Hermitian matrices.

The interesting metrics on such spaces are the ones that are invariant under all the symmetries, in particular the invariant Finsler metrics. When the symmetric space is non-compact, as in the example just referred to, it is profitable to study the horofunction boundary of such metrics.

An important technique in trying to understand symmetric spaces is to look at their *flats*. These are subspaces that are, as their name suggests, flat in some sense. Because of the abundance of symmetries, there are many flats; indeed, every pair of points lies in a flat. Furthermore, given any two flats, there is a symmetry taking one to the other, and so they are all alike. It turns out that the restriction of an invariant Finsler metric to a single flat determines the metric everywhere, and gives the flat the geometry of a normed space.

Symmetric spaces can be compactified by means of the Satake compactification. In fact, there are several such compactifications, one associated to each irreducible faithful representation of the invariance group of the space. In [41], we show that each Satake compactification can be constructed as a horofunction compactification by choosing an appropriate invariant Finsler metric. In fact, the metrics we construct have polyhedral balls on the flat.

An important step in the proof is to show that the closure of a flat in the horofunction compactification of the symmetric space is the same as the horofunction compactification of the flat viewed as a metric space in its own right. This is not true for every metric space, since in general one might not be able to distinguish horofunctions by looking at a subspace.

7.2.3. The set of minimal upper bounds of two matrices in the Loewner order

Participant: Nikolas Stott.

A classical theorem of Kadison shows that the space of symmetric matrices equipped with the Loewner order is an anti-lattice, meaning that two matrices have a least upper bound if and only if they are comparable. In [24], we refined this theorem by characterizing the set of minimal upper bounds: we showed that it is homeomorphic to the quotient space $O(p) \setminus O(p,q)/O(q)$, where $O(p,q)$ denotes the orthogonal group associated to the quadratic form with signature (p,q), and $O(p)$ denotes the standard pth orthogonal group.

7.2.4. Generalization of the Hellinger distance

Participant: Stéphane Gaubert.

In [64] (joint work with Rajendra Bhatia of Ashoka University and Tanvi Jain, Indian Statistic Institute, New Delhi), we study some generalizations of the Hellinger distance to the space of positive definite matrices.

7.2.5. Spectral inequalities for nonnegative tensors and their tropical analogues

Participant: Stéphane Gaubert.

In [39] (joint work with Shmuel Friedland, University of Illinois at Chicago) we extend some characterizations and inequalities for the eigenvalues of nonnegative matrices, such as Donsker-Varadhan, Friedland-Karlin, Karlin-Ost inequalities, to nonnegative tensors. These inequalities are related to a correspondence between nonnegative tensors and ergodic control: the logarithm of the spectral radius of a tensor is given by the value of an ergodic problem in which instantaneous payments are given by a relative entropy. Some of these inequalities involve the tropical spectral radius, a limit of the spectral radius which we characterize combinatorially as the value of an ergodic Markov decision process.

7.3. Tropical algebra and convex geometry

7.3.1. Formalizing convex polyhedra in Coq

Participants: Xavier Allamigeon, Ricardo Katz [Conicet, Argentine].

In [20], we have made the first steps of a formalization of the theory of convex polyhedra in the proof assistant Coq. The originality of our approach lies in the fact that our formalization is carried out in an effective way, in the sense that the basic predicates over polyhedra (emptiness, boundedness, membership, etc) are defined by means of Coq programs. All these predicates are then proven to correspond to the usual logical statements. The latter take the form of the existence of certificates: for instance, the emptiness of a polyhedron is shown to be equivalent to the existence of a certificate *a la* Farkas. This equivalence between Boolean predicates and formulas living in the kind `Prop` is implemented by using the boolean reflection methodology, and the supporting tools provided by the Mathematical Components library and its tactic language. The benefit of the effective nature of our approach is demonstrated by the fact that we easily arrive at the proof of important results on polyhedra, such as several versions of Farkas Lemma, duality theorem of linear programming, separation from convex hulls, Minkowski Theorem, etc.

Our effective approach is made possible by implementing the simplex method inside Coq, and proving its correctness and termination. Two difficulties need to be overcome to formalize it. On the one hand, we need to deal with its termination. More precisely, the simplex method iterates over the so-called bases. Its termination depends on the specification of a pivoting rule, whose aim is to determine, at each iteration, the next basis. In this work, we have focused on proving that the lexicographic rule ensures termination. On the other hand, the simplex method is actually composed of two parts. The part that we previously described, called Phase II, requires an initial basis to start with. Finding such a basis is the purpose of Phase I. It consists in building an extended problem (having a trivial initial basis), and applying to it Phase II. Both phases need to be formalized to obtain a fully functional algorithm.

7.3.2. Tropical totally positive matrices

Participant: Stéphane Gaubert.

In [22] (joint work with Adi Niv) we investigate the tropical analogues of totally positive and totally non-negative matrices, i.e, the images by the valuation of the corresponding classes of matrices over a non-archimedean field. We show in particular that tropical totally positive matrices essentially coincide with the Monge matrices (defined by the positivity of 2×2 tropical minors), arising in optimal transport, and compare the set of tropical totally positive matrices with the tropicalization of the totally positive Grassmannian.

7.3.3. Tropical compound matrix identities

Participants: Marianne Akian, Stéphane Gaubert.

A number of polynomial identities in tropical semirings can be derived from their classical analogues by application of a transfer principle [49], [51]. In [16], joint with Adi Niv, we prove identities on compound matrices in extended tropical semirings, which cannot be obtained by transfer principles, but are rather obtained by combinatorial methods. Such identities include analogues to properties of conjugate matrices, powers of matrices and $\mathrm{adj}\,(A)\,\det\,(A)^{-1}$, all of which have implications on the eigenvalues of the corresponding matrices. A tropical Sylvester-Franke identity is provided as well.

7.3.4. Group algebra in characteristic one and invariant distances over finite groups

Participant: Stéphane Gaubert.

In [21] (joint work with Dominique Castella), we investigated a tropical analogue of group algebras. We studied tropical characters and related them to invariant distances over groups.

7.3.5. Volume and integer points of tropical polytopes

Participant: Stéphane Gaubert.

We investigate in [40] (joint work with Marie McCaig) the volume of tropical polytopes, as well as the number of integer points contained in integer polytopes. We proved that even approximating these values for a tropical polytope given by its vertices is hard, with no approximation algorithm with factor $2^{\mathrm{poly}(m,n)}$ existing unless $\mathrm{P} = \mathrm{NP}$.

7.4. Tropical methods applied to optimization, perturbation theory and matrix analysis

7.4.1. Tropicalization of the central path and application to the complexity of interior point methods

Participants: Xavier Allamigeon, Stéphane Gaubert.

This work is in collaboration with Pascal Benchimol (EDF Labs) and Michael Joswig (TU Berlin).

In optimization, path-following interior point methods are driven to an optimal solution along a trajectory called the central path. The *central path* of a linear program $\mathrm{LP}(A, b, c) \equiv \min\{c \cdot x \mid Ax \leq b,\ x \geq 0\}$ is defined as the set of the optimal solutions (x^μ, w^μ) of the barrier problems:

$$\begin{array}{ll} \text{minimize} & c \cdot x - \mu(\sum_{j=1}^{n} \log x_j + \sum_{i=1}^{m} \log w_i) \\ \text{subject to} & Ax + w = b,\ x > 0,\ w > 0 \end{array}$$

While the complexity of interior point methods is known to be polynomial, an important question is to study the number of iterations which are performed by interior point methods, in particular whether it can be bounded by a polynomial in the dimension (mn) of the problem. This is motivated by Smale 9th problem [128], on the existence of a strongly polynomial complexity algorithm for linear programming. So far, this question has been essentially addressed though the study of the curvature of the central path, which measures how far a path differs from a straight line, see [77], [76], [79], [78]. In particular, by analogy with the classical Hirsch conjecture, Deza, Terlaky and Zinchencko [78] proposed the “continuous analogue of the Hirsch conjecture”, which says that the total curvature of the central path is linearly bounded in the number m of constraints.

In a work of X. Allamigeon, P. Benchimol, S. Gaubert, and M. Joswig [17], we prove that primal-dual log-barrier interior point methods are not strongly polynomial, by constructing a family of linear programs with $3r + 1$ inequalities in dimension $2r$ for which the number of iterations performed is in $\Omega(2^r)$. The total curvature of the central path of these linear programs is also exponential in r, disproving the continuous analogue of the Hirsch conjecture.

Our method is to tropicalize the central path in linear programming. The tropical central path is the piecewise-linear limit of the central paths of parameterized families of classical linear programs viewed through logarithmic glasses. We give an explicit geometric characterization of the tropical central path, as a tropical analogue of the barycenter of a sublevel set of the feasible set induced by the duality gap. We study the convergence properties of the classical central path to the tropical one. This allows us to show that that the number of iterations performed by interior point methods is bounded from below by the number of tropical segments constituting the tropical central path.

7.4.2. Tropicalization of semidefinite programming and its relation with stochastic games

Participants: Xavier Allamigeon, Stéphane Gaubert, Mateusz Skomra.

Semidefinite programming consists in optimizing a linear function over a spectrahedron. The latter is a subset of $\mathbb{R}^n$ defined by linear matrix inequalities, i.e., a set of the form

$$\left\{x \in \mathbb{R}^n : Q^{(0)} + x_1 Q^{(1)} + \cdots + x_n Q^{(n)} \succeq 0\right\}$$

where the $Q^{(k)}$ are symmetric matrices of order m, and $\succeq$ denotes the Loewner order on the space of symmetric matrices. By definition, $X \succeq Y$ if and only if $X - Y$ is positive semidefinite.

Semidefinite programming is a fundamental tool in convex optimization. It is used to solve various applications from engineering sciences, and also to obtain approximate solutions or bounds for hard problems arising in combinatorial optimization and semialgebraic optimization.

A general issue in computational optimization is to develop combinatorial algorithms for semidefinite programming. Indeed, semidefinite programs are usually solved via interior point methods. However, the latter provide an approximate solution in a polynomial number of iterations, provided that a strictly feasible initial solution. Semidefinite programming becomes a much harder matter if one requires an exact solution. The feasibility problem belongs to $\mathsf{NP}_{\mathbb{R}} \cap \mathsf{coNP}_{\mathbb{R}}$, where the subscript $\mathbb{R}$ refers to the BSS model of computation. It is not known to be in NP in the bit model.

We address semidefinite programming in the case where the field $\mathbb{R}$ is replaced by a nonarchimedean field, like the field of Puiseux series. In this case, methods from tropical geometry can be applied and are expected to allow one, in generic situations, to reduce semialgebraic problems to combinatorial problems, involving only the nonarchimedean valuations (leading exponents) of the coefficients of the input.

To this purpose, we first study tropical spectrahedra, which are defined as the images by the valuation of nonarchimedean spectrahedra. We establish that they are closed semilinear sets, and that, under a genericity condition, they are described by explicit inequalities expressing the nonnegativity of tropical minors of order 1 and 2. These results are presented in the preprint [60], with further results in the PhD thesis [13].

We show in [18] that the feasibility problem for a generic tropical spectrahedron is equivalent to solving a stochastic mean payoff game (with perfect information). The complexity of these games is a long-standing open problem. They are not known to be polynomial, however they belong to the class $\mathsf{NP} \cap \mathsf{coNP}$, and they can be solved efficiently in practice. This allows to apply stochastic game algorithms to solve nonarchimedean semidefinite feasibility problems. We obtain in this way both theoretical bounds and a practicable method which solves some large scale instances.

A long-standing problem is to characterize the convex semialgebraic sets that are SDP representable, meaning that they can be represented as the image of a spectrahedron by a (linear) projector. Helton and Nie conjectured that every convex semialgebraic set over the field of real numbers are SDP representable. Recently, [124] disproved this conjecture. In [19], we show, however, that the following result, which may be thought of as a tropical analogue of this conjecture, is true: over a real closed nonarchimedean field of Puiseux series, the convex semialgebraic sets and the projections of spectrahedra have precisely the same images by the nonarchimedean valuation. The proof relies on game theory methods and on our previous results [60] and [18].

In [27] and [13], we exploit the tropical geometry approach to introduce a condition number for stochastic mean payoff games (with perfect information). This condition number is defined as the maximal radius of a ball in Hilbert's projective metric, contained in a primal or dual feasible set. We show that the convergence time of value iteration is governed by this condition number, and derive fixed parameter tractability results.

7.4.3. Tropical polynomial systems and colorful interior of convex bodies

Participants: Marianne Akian, Marin Boyet, Xavier Allamigeon, Stéphane Gaubert.

The starting PhD thesis work of Marin Boyet, deals with the solution of tropical polynomial systems, with motivations from call center performance evaluation (see Section 7.6.1). We introduced a notion of colorful interior of a family of convex bodies, and showed that the solution of such a polynomial system reduces to linear programming if one knows a vector in the colorful interior of an associated family of Newton polytopes. Further properties of colorful interiors are currently investigated.

7.5. Tropical algebra, number theory and directed algebraic topology

7.5.1. An arithmetic site of Connes-Consani type for number fields with narrow class number 1

Participant: Aurélien Sagnier.

In 1995, A. Connes ([71]) gave a spectral interpretation of the zeroes of the Riemann zeta function involving the action of $\mathbb{R}_+^*$ on the sector $X = \mathbb{Q}_+^\times \setminus \mathbb{A}_\mathbb{Q}/\widehat{\mathbb{Z}}^\times$ of the adele class space $\mathbb{A}_\mathbb{Q}/\mathbb{Q}^*$ of the field of rational numbers. In [72], [74], the action of $\mathbb{R}_+^*$ on this sector X was shown to have a natural interpretation in algebraic geometry. This interpretation requires the use of topos theory as well as of the key ingredient of characteristic one namely the semifield $\mathbb{R}_{\max}$ familiar in tropical geometry. The automorphism group of this semifield is naturally isomorphic to $\mathbb{R}_+^*$ and plays the role of the Frobenius. As it turns out, its action on the points of a natural semiringed topos corresponds canonically to the above action on X. This semiringed topos is called the arithmetic site. In my PhD, I extended the construction of the arithmetic site, replacing the field of rational numbers by certain number fields. I considered the simplest complex case, namely that of imaginary quadratic fields on which we assume that the units are not reduced to ± 1 that is when K is either $\mathbb{Q}(\imath)$ or $\mathbb{Q}(\imath\sqrt{3})$. These results are presented in the submitted article [121]. In a further work, developed this year, I extended this construction, dealing now with number fields K with narrow class number 1. In fact, if we denote $\mathcal{U}_{\mathcal{O}_K}^+$ the totally positive units of $\mathcal{O}_K$, $\mathcal{O}_K^+$ the totally positive integers, $\widehat{\mathcal{O}_K^+}$ the topos of sets with a multiplicative action of totally positive integers of $\mathcal{O}_K^+$, $D_K = \text{Conv}(\mathcal{U}_{\mathcal{O}_K}^+)$ and $\mathcal{C}_{\mathcal{O}_K} = \text{Semiring}(\{\lambda D_K/\lambda \in \mathcal{O}_K^+\}) \cup \{\varnothing, \{0\}\}$, we consider the semiringed topos $(\widehat{\mathcal{O}_K^+}, (\mathcal{C}_{\mathcal{O}_K}, \text{Conv}(\bullet \cup \bullet), +))$ and show for it similar properties as the one shown in my PhD thesis for the arithmetic sites associated to imaginary quadratic fields with class number 1 by adapting to this case the technics used in my PhD thesis, Shintani units theorem, and some remarks A.Connes made on my PhD which appear as an appendix of the article [121]. Here again tropical algebra play a crucial role in the geometrical constructions.

7.5.2. Tropical tensor products

Participants: Stéphane Gaubert, Aurélien Sagnier.

Tensors products of modules over semifields of characteristic one, like the Boolean or tropical semifields, have appeared recently, with motivations from arithmetics, in work by Connes and Consani, towards an intersection theory for arithmetic and scaling sites (spaces they have built and which are closely related to the zeroes of the Riemann zeta function). Algebraic and topological tropical tensors products were constructed in a different way by Litvinov and collaborators: here, tropical tensors are sums of "rank one" expressions, similar to the ones used in the approximation of large data sets or of functions of many parameters. We show that the canonical notion of tropical tensor product, defined in terms of the usual universal problem, differs from the definition arising from approximation theory, but that the latter can be recovered from the former by a certain "reduction" operation. We illustrate these results by computing several basic examples of categorical tensors products, including spaces of convex sets and functions.

7.5.3. Directed topological complexity and control

Participant: Aurélien Sagnier.

This is a joint work with Michael Farber and Eric Goubault.

The view we are taking here is that of topological complexity, as defined in [83], adapted to directed topological spaces.

Let us briefly motivate the interest of a directed topological complexity notion. It has been observed that the very important planification problem in robotics boils down to, mathematically speaking, finding a section to the path space fibration $\chi \ : \ PX = X^I \to X \times X$ with $\chi(p) = (p(0), p(1))$. If this section is continuous, then the complexity is the lowest possible (equal to one), otherwise, the minimal number of discontinuities that would encode such a section would be what is called the topological complexity of X. This topological complexity is both understandable algorithmically, and topologically, e.g. as s having a continuous section is equivalent to X being contractible. More generally speaking, the topological complexity is defined as the Schwartz genus of the path space fibration, i.e. is the minimal cardinal of partitions of $X \times X$ into “nice” subspaces F_i such that $s_{F_i} \ : \ F_i \to PX$ is continuous.

This definition perfectly fits the planification problem in robotics where there are no constraints on the actual control that can be applied to the physical apparatus that is supposed to be moved from point a to point b. In many applications, a physical apparatus may have dynamics that can be described as an ordinary differential equation in the state variables $x \in \mathbb{R}^n$ and in time t, parameterized by control parameters $u \in \mathbb{R}^p$, $\dot{x}(t) = f(t, x(t))$. These parameters are generally bounded within some set U, and, not knowing the precise control law (i.e. parameters u as a function of time t) to be applied, the way the controlled system can evolve is as one of the solutions of the differential inclusion $\dot{x}(t) \in F(t, x(t))$ where $F(t, x(t))$ is the set of all $f(t, x(t), u)$ with $u \in U$. Under some classical conditions, this differential inclusion can be proven to have solutions on at least a small interval of time, but we will not discuss this further here. Under the same conditions, the set of solutions of this differential inclusion naturally generates a dspace (a very general structure of directed space, where a preferred subset of paths is singled out, called directed paths, see e.g. [97]). Now, the planification problem in the presence of control constraints equates to finding sections to the analogues to the path space fibration (That would most probably not qualify for being called a fibration in the directed setting) taking a dipath to its end points. This notion is developed in this article, and we introduce a notion of directed homotopy equivalence that has precisely, and in a certain non technical sense, minimally, the right properties with respect to this directed version of topological complexity.

This notion of directed topological complexity also has applications in informatics where a directed space can be used to model the space of all possible executions of a concurrent process (ie when several running programs must share common limited ressources).

In a recent prepublication [84], after defining the notion of directed topological complexity,this invariant (directed topological complexity) is studied for directed spheres and directed graphs.

7.6. Applications

7.6.1. Performance evaluation of an emergency call center

Participants: Xavier Allamigeon, Stéphane Gaubert.

Since 2014, we have been collaborating with Préfecture de Police (Régis Reboul and LcL Stéphane Raclot), more specifically with Brigade de Sapeurs de Pompiers de Paris (BSPP) and Direction de Sécurité de Proximité de l'agglomération parisienne (DSPAP), on the performance evaluation of the new organization to handle emergency calls to firemen and policemen in the Paris area. We developed analytical models, based on Petri nets with priorities, and fluid limits, see [55], [56], [65]. In 2018, we performed specific case studies, with several students of École polytechnique: Laetitia de Coudenhove, Julie Poulet, Céline Moucer and Julia Escribe.

7.6.2. Tropical models of fire propagation in urban areas

Participants: Stéphane Gaubert, Daniel Jones.

As part of the team work in the ANR project Democrite, we developed a model of fire propagation in urban areas, involving a deterministic analogue of first passage percolation. We showed that the fire tends to propagate according to a polyhedral shape, and derived metric limit theorems, exploiting discrete convexity results à la Shapley-Folkman. We validated this approach by simulations, on the fire following Kobe earthquake in 1995. The polyhedral shape is also apparent in historical fires, like the great fire of London (1666). These results are announced in [28].

7.6.3. Smart Data Pricing

Participants: Marianne Akian, Jean-Bernard Eytard, Stéphane Gaubert.

This work is in collaboration with Mustapha Bouhtou (Orange Labs) and with Gleb Koshevoy (Russian academy of Science).

The PhD work [81] of Jean-Bernard Eytard concerns the optimal pricing of data trafic in mobile networks. We developed a bilevel programming approach, allowing to an operator to balance the load in the network through price incentives. We showed that a subclass of bilevel programs can be solved in polynomial time, by combining methods of tropical geometry and of discrete convexity. This work is presented in [80] and also in [34]. In a followup work, presented in [81], we managed to extend these results to wider classes of bilevel problems, and to relate them to competitive equilibria problems.

7.6.4. Game theory models of decentralized mechanisms of pricing of the smart grid

Participants: Stéphane Gaubert, Paulin Jacquot.

This work is in collaboration with Nadia Oudjane, Olivier Beaude and Cheng Wan (EDF Labs).

The PhD work of Paulin Jacquot concerns the application of game theory techniques to pricing of energy. We are developing a game theory framework for demand side management in the smart grid, in which users have movable demands (like charging an electric vehicle). We compared in particular the daily and hourly billing mechanisms. The latter, albeit more complex to analyse, has a merit as it incitates the user to move his or her consumption at off peak hours. We showed the Nash equilibrium is unique, under some assumptions, and gave theoretical bounds of the price of anarchy of the game with a hourly billing, showing this mechanism remains efficient while being more “fair” than the daily billing. We proposed and tested decentralized algorithms to compute the Nash equilibrium. These contributions are presented in [102], [23].

Another work, by Paulin Jacquot and Cheng Wan, deals with limit theorems for atomic games with a large number of players [30], [42].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- Yield management methods applied to the pricing of data trafic in mobile networks. CRE (research contract) with Orange Labs (Orange Labs partner: Mustapha Bouhtou).
- Decentralized mechanisms of operation of power systems: equilibria and efficiency. Collaboration with Nadia Oudjane and Olivier Beaude from EDF-labs, with the PhD work of Paulin Jacquot (CIFRE PhD), supervised by Stéphane Gaubert.
- Stochastic optimization of multiple flexibilities and energies in micro-grids, collaboration with Wim Van Ackooij, from EDF labs, with the PhD work of Maxime Grangereau (CIFRE PhD), supervised by Emmanuel Gobet (CMAP) and cosupervised by Stéphane Gaubert.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

- Projet ANR MALTHY (Méthodes ALgébriques pour la vérification de modèles Temporisés et HYbrides), responsable T. Dang. Partenaires : Verimag, CEA LIST, Inria Rennes, Inria Saclay, VISEO/Object Direct.
- Projet ANR DEMOCRITE ("DEmonstrateur d’un MOteur de Couverture des Risques sur un TErritoire), responsable Emmanuel Lapébie (CEA). Partenaires : CEA-GRAMAT, BSPP, Inria Saclay (Maxplus), Institut PPRIME - UPR3346 (CNRS, Univ. Poitiers, ISAE-ENSMA), IPSIS, SYSTEL, ARMINES-E.M. Alès-ISR, CERDACC (Univ. de Haute-Alsace).

- Projet ANR JCJC CAPPS ("Combinatorial Analysis of Polytopes and Polyhedral Subdivisions"), responsable Arnau Padrol (IMJ-PRG, Sorbonne Université). Partenaires : IMJ-PRG (Sorbonne Université), Inria Saclay (Tropical), LIGM (Université Paris-Est Marne-la-Vallée), LIF (Université Aix-Marseille), CERMICS (École Nationale des Ponts et Chaussées), LIX (École Polytechnique).

9.1.2. Programme Gaspard Monge pour l'Optimisation

- Projet intitulé "Méthodes tropicales pour l'optimisation", responsable X. Allamigeon, faisant intervenir M. Akian, V. Boeuf, S. Gaubert, A.Hochart, R. Katz, et M. Skomra.

9.2. European Initiatives

9.2.1. Collaborations with Major European Organizations

- Partner: Michael Joswig, TU-Berlin.
- Topic : Tropical geometry.

9.3. International Initiatives

9.3.1. Inria International Partners

9.3.1.1. Informal International Partners

- Collaboration with Ricardo D. Katz, CIFASIS-CONICET, Rosario (Argentina). Research invitation at CMAP during 2 months.
- Collaboration with Shmuel Friedland, University of Illinois at Chicago (invitation or Stéphane Gaubert at Chicago of one week in May 2018).
- Collaboration with Alejandro Jofre, CMM, University of Chili, Santiago: invitation of Paulin Jacquot of two months (May-June) 2018.
- Collaboration with Zheng Qu, Math. Department, Hongk Kong University.

9.3.2. Participation in Other International Programs

- Collaboration with Gleb Koshevoy, Poncelet Laboratory, Moscow (research invitation of Gleb Koshevoy at CMAP during one week).
- Collaboration with Aris Daniilidis, from CMM, University of Chile, Santiago.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

- Aris Daniilidis, from CMM, University of Chile, Santiago, Sept-Dec 2018. Invited to École polytechnique within the Gaspard Monge Visiting Professor Program (GMVP) of École polytechnique, with the support of Fondation de l'École polytechnique.
- Roberto Bobadilla Solari, invited PhD student, University of Chile, Santiago, from Sep 2018 until Nov 2018, funded by Inria, associated to the visit of Aris Daniilidis.
- Gonzalo Flores Garcia, invited PhD student, University of Chile, Santiago, from Sep 2018 until Dec 2018, associated to the visit of Aris Daniilidis.
- Sebastian Tapia Garcia, invited PhD student, University of Chile, Santiago, funded by GMVP, associated to the visit of Aris Daniilidis.
- Francisco Javier Antonio Venegas Martinez, invited PhD student, University of Chile, Santiago, from Sep 2018 until Nov 2018, funded by Inria, associated to the visit of Aris Daniilidis.
- Rajendra Bhatia (Ashoka University, India), June 2018.
- Gleb Koshevoy (Russian Academy of Siences), Nov 2018.

9.4.1.1. Internships

- Raphael Pellegrin (Imperial College, London), research summer internship, on tropical positivstellensätze.
- Marin Boyet (École Nationale Supérieure des Mines de Paris), research internship, colorful interior of convex bodies and solvability of tropical polynomial systems.

9.4.2. Visits to International Teams

9.4.2.1. Research Stays Abroad

- Marianne Akian, Institut Mittag Leffler, Jan 15-March 3, 2018.
- Stéphane Gaubert, Institut Mittag Leffler, Jan 15-March 3, 2018.
- Stéphane Gaubert, joint invitation by the Statistics Department of University of Chicago (Lek Heng Lim) and the Maths and Computer Science Department of University of Illinois at Chicago (Shmuel Friedland), May 20-26, 2018.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- Stéphane Gaubert is the coordinator of the Gaspard Monge Program for Optimization, Operations Research and their interactions with data sciences (PGMO), a corporate sponsorhip program, operated by Fondation Mathématique Jacques Hadamard, supported by Criteo, EDF, Orange and Thales, see http://www.fondation-hadamard.fr/fr/pgmo/.

10.1.1.2. Member of the Organizing Committees

- S. Gaubert co-organizes the "Séminaire Parisien d'Optimisation" at Institut Henri Poincaré. https://sites.google.com/site/spoihp/.

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

- S. Gaubert, Chair of the PGMO days, EDF Labs Paris Saclay, Nov 20-21, 2018. http://www.fondation-hadamard.fr/fr/pgmo/pgmodays
- S. Gaubert, coorganizer, with A. Daniilidis and S. Tapia, of the workshop "Dynamical Aspects in Variational Analysis", École polytechnique, Dec. 14th, 2018. http://www.cmap.polytechnique.fr/~gaubert/VariationalAnalysisWorkshop/index.html

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- S. Gaubert is member of the editorial committee of the collection Mathématiques et Applications, SMAI and Springer.
- S. Gaubert is associate editor of Linear and Multilinear Algebra.
- S. Gaubert is associate editor of RAIRO Operations research.

10.1.4. Invited Talks

- M. Akian
  - Majorization inequalities for valuations of eigenvalues, Mittag-Leffler institute, February 2018.

- X. Allamigeon
  – Log-barrier interior point methods are not strongly polynomial, Workshop on Tropical Algebra and Applications, Mittag-Leffler Institute, Stockholm, Sweden, January 2018
  – Log-barrier interior point methods are not strongly polynomial, Workshop on Algebraic and geometric aspects of semidefinite programming at ISMP, Bordeaux, France, June 2018
  – Performance evaluation of an emergency call center: tropical polynomial systems applied to timed Petri nets, Workshop on Tropical Mathematics & Optimisation for Railways, Birmingham University, UK, June 2018
  – First steps in the formalization of convex polyhedra in Coq, International Conference on Mathematical Software, Notre-Dame University, USA, July 2018
- S. Gaubert
  – Nonarchimedean convex programming and its relation to mean payoff games, invited plenary talk at Highlights of Logic, Automata, and Games, Berlin, September 2018.
  – Condition numbers in nonarchimedean semidefinite programming and their relation with stochastic mean payoff games, Mittag-Leffler institute, February 2018.
  – Inequalities for the spectral radii and spectral norms of nonnegative tensors, ISMP, Bordeaux, Invited session, July 2018.
  – Nonarchimedean Degenerations of Convex Programming – In Which Tropical Geometry Shows That Log-Barrier Interior Point Methods Are Not Strongly Polynomial, Computational and Applied Mathematics Colloquium, The University of Chicago, May 2018.

10.1.5. Leadership within the Scientific Community

- See Section 10.1.1.1 (coordination of PGMO).

10.1.6. Research Administration

- M. Akian :
  – Member of the “comité de liaison SMAI-MODE” since June 2015.
  – Member of the laboratory council of CMAP.
- S. Gaubert :
  – Chairman of the Gaspard Monge Program for Optimization, Operations Research and their interactions with data sciences (PGMO), see 10.1.1.1 for details.
  – Member of the scientific council of CMAP.
- X. Allamigeon:
  – Member of the scientific committee of Inria Saclay – Ile-de-France.
  – Member of the Applied Mathematics Department committee at Ecole Polytechnique.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- M. Akian
  – Course “Markov decision processes: dynamic programming and applications” joint between (3rd year of) ENSTA and M2 “Mathématiques et Applications”, U. Paris Saclay, “Optimization”, and shared with Jean-Philippe Chancelier (ENPC), 15 hours each.
- X. Allamigeon

  - Petites classes et encadrement d'enseignements d'approfondissement de Recherche Opérationnelle en troisième année à l'École Polytechnique (programme d'approfondissement de Mathématiques Appliquées) (niveau M1).
  - Cours du M2 "Optimisation" de l'Université Paris Saclay, cours partagé avec Manuel Ruiz (RTE) et Dominique Quadri (LRI, Université Paris Sud).
  - Co-responsabilité du programme d'approfondissement en mathématiques appliquées (troisième année) à l'École Polytechnique.
- J.B. Eytard
  - Cours de niveau L1-L2 à l'IUT d'Informatique d'Orsay (Univ. Paris-Sud), dans le cadre d'un monitorat (64h) (théorie des graphes, recherche opérationnelle, modélisation mathématique).
- S. Gaubert
  - Course "Systèmes à Événements Discrets", option MAREVA, ENSMP.
  - Course "Algèbre tropicale pour le contrôle optimal et les jeux" of "Contrôle, Optimisation et Calcul des Variations" (COCV) of M2 "Mathématiques et Applications" of Paris 6 University and École Polytechnique.
  - Lecture of Operations Research, third year of École Polytechnique. The lectures notes were published as a book [66].
- M. Skomra
  - TD de mathématiques à l'UPMC.

10.2.2. Supervision

- PhD : Eric Fodjo, registered at École Polytechnique, since October 2013, thesis supervisor: Marianne Akian, defended on Jul 13 2017.
- PhD : Mateusz Skomra, registered at Univ. Paris Saclay since October 2015, thesis supervisor: Stéphane Gaubert, cosupervision: Xavier Allamigeon, defended on Dec. 5, 2018.
- PhD : Jean-Bernard Eytard, registered at Univ. Paris Saclay since October 2015, thesis supervisor: Stéphane Gaubert, cosupervision: Marianne Akian, Mustapha Bouhtou, defended on Nov. 12, 2018.
- PhD in progress: Paulin Jacquot, registered at Univ. Paris Saclay since November 2016, thesis supervisor: Stéphane Gaubert, cosupervision: Nadia Oujdane, Olivier Beaude (EDF).
- PhD in progress: Benoît Tran, registered at Univ Paris-Est Marne La Vallée, since September 2017, thesis supervisor: Jean-Philippe Chancelier (ENPC), cosupervision: Marianne Akian.
- PhD in progress: Maxime Grangereau, registered at Univ. Paris Saclay since Jan 2018, thesis supervisor: Emanuel Gobet, cosupervision: Stéphane Gaubert.
- PhD in progress: Omar Saadi, registered at Univ. Paris Saclay since October 2018, thesis supervisor: Stéphane Gaubert, cosupervision: Marianne Akian.
- PhD in progress: Marin Boyet, registered at Univ. Paris Saclay since October 2018, thesis supervisor: Stéphane Gaubert, cosupervision: Xavier Allamigeon.

10.2.3. Juries

- M. Akian
  - Jury of the 2018 competition for CR2 positions of Inria Saclay–Île-de-France.
  - Jury of the PhD thesis of E. Fodjo (Ecole Polytechnique, thesis supervisor, July 2018).
  - Jury of the PhD thesis of H. Gerard (ENPC, reviewer, Oct. 2018).
  - Jury of the PhD thesis of J.B. Eytard (Ecole Polytechnique, examiner, Nov. 2018).
- X. Allamigeon

  - Comité de suivi doctoral de D. Rouhling (Inria Sophia, Mai 2018)
  - Jury of the PhD thesis of M. Skomra (Ecole Polytechnique, examiner, Dec. 2018).
- S. Gaubert
  - Jury of the HdR of Welington de Oliveira (Sorbonne Universités –Paris I– examiner, Déc. 2018)
  - Jury of the PhD thesis of J.B. Eytard (Ecole Polytechnique, examiner, Nov. 2018).
  - Jury of the PhD thesis of T. Roget (Ecole Polytechnique, examiner, Nov. 2018).
  - Jury of the PhD thesis of M. Skomra (Ecole Polytechnique, examiner, Dec. 2018).

10.3. Conferences, Seminars

- M. Akian
  - The operator approach to entropy games, International workshop on game theory, IHP, June 2018, Paris.
  - Majorization inequalities for valuations of eigenvalues, ANR/DFG PRCI Project Kick-off Meeting, July 2018, Bonn.
  - Tropical geometry, Optimal Control and Mean-payoff Games, Dagstuhl Seminar "Shape Analysis: Euclidean, Discrete and Algebraic Geometric Methods", Oct. 2018, Dagstuhl.
- X. Allamigeon
  - Tropical Linear Optimization, Polytopes à Paris, Paris, France, January 2018.
  - A formalization of convex polyhedra based on the simplex method, Séminaire de Géométrie Algorithmique et Combinatoire, Paris, France, March 2018.
  - Log-barrier interior point methods are not strongly polynomial, Journée de rentrée du CMAP, Palaiseau, France, October 2018.
  - First steps in the formalization of convex polyhedra in Coq, Seminar on Solvers Principles and Architectures, Rennes, France, October 2018.
- M. Boyet
  - The shadow vertex algorithm solves colorful one-versus-all tropical polynomial systems, PGMO Days, Nov. 21, 2018, Palaiseau.
- J.B. Eytard
  - A tropical approach to bilevel programming and a comparison with a competitive equilibrium problem, ROADEF, Feb. 2018, Université Bretagne Sud, Lorient.
  - Tropical geometry applied to bilevel programming and an application to price incentives in telecom networks, JFRO (journées franciliennes de recherche opérationnelle) sur le thème "optimisation bi-niveaux", March 2018, CNAM, Paris.
  - Tropical geometry and discrete convexity applied to bilevel programming, IWOBIP 18, June 2018, Lille.
  - Tropical geometry applied to bilevel programming, ISMP, July 2018, Bordeaux.
  - How to use tropical geometry to solve bilevel programming problems ?, PGMO Days, Nov. 21, 2018, Palaiseau.
- S. Gaubert
  - Tropical spectrahedra and their relation with mean payoff games, University of Illinois at Chicago, May 2018.
  - Tropical spectrahedra and their relation with mean payoff games, International workshop on game theory, IHP, June 2018.

  - A convergent hierarchy of non-linear eigenproblems to compute the joint spectral radius of nonnegative matrices, 23rd International Symposium on Mathematical Theory of Networks and Systems (MTNS), Hong-Kong, Jul. 2018.
- P. Jacquot
  - Fast Computation of Equilibria in Splittable Routing Games: Application to Electricity Demand Response, SMAI Mode, Autrans, March 2018
  - Analysis of a Routing Game Model for Demand Side Management, ISMP, Bordeaux, July 2018.
  - Routing Game on Parallel Networks: the Convergence of Atomic to Nonatomic, CDC, Miami, December 2018
- M. Skomra
  - The condition number of stochastic mean payoff games, ISMP, Bordeaux, July 2018.
  - Condition Numbers of Stochastic Mean Payoff Games and What They Say about Nonarchimedean Semidefinite Programming, 23rd International Symposium on Mathematical Theory of Networks and Systems (MTNS), Hong-Kong, July, 2018.
- A. Sagnier
  - Talk at Ohio State University on «An arithmetic site of Connes-Consani type for $\mathbb{Z}[\imath]$», January 2018.
  - Talk at Rutgers University on «An arithmetic site of Connes-Consani type for $\mathbb{Z}[\imath]$», January 2018
  - Talk at Johns Hopkins University on «An arithmetic site of Connes-Consani type for $\mathbb{Z}[\imath]$», January 2018
  - Short communication at Toposes in Como on «An arithmetic site of Connes-Consani type for $\mathbb{Z}[\sqrt{2}]$», Università degli Studi dell'Insubria, Como, June 2018
  - Short communication at the International Congress of Mathematicians, ICM 2018 on «An arithmetic site of Connes-Consani type for $\mathbb{Z}[\imath]$» , August 2018
  - Short communication at the Tensors conference on «Tropical tensor products» Polytecnico di Torino, Turin, September 2018.
  - Talk at University of Oslo on «An arithmetic site of Connes-Consani type for for number fields with narrow class number 1», November 2018.
- B. Tran
  - A Stochastic Min-plus Algorithm for Deterministic Optimal Control, ISMP, July 2018, Bordeaux.
- C. Walsh
  - Seminar, Université de Nantes, December 14, 2018. Title of the talk: "The horofunction boundary of Teichmüller space".

11. Bibliography

Major publications by the team in recent years

[1] M. AKIAN, S. GAUBERT, R. BAPAT.*Non-archimedean valuations of eigenvalues of matrix polynomials*, in "Linear Algebra and its Applications", June 2016, vol. 498, p. 592–627, Also arXiv:1601.00438 [*DOI :* 10.1016/J.LAA.2016.02.036], https://hal.inria.fr/hal-01251803

[2] M. AKIAN, S. GAUBERT, A. GUTERMAN.*Tropical polyhedra are equivalent to mean payoff games*, in "Internat. J. Algebra Comput.", 2012, vol. 22, n^{o} 1, 1250001, 43, http://dx.doi.org/10.1142/S0218196711006674

[3] M. AKIAN, S. GAUBERT, R. NUSSBAUM.*Uniqueness of the fixed point of nonexpansive semidifferentiable maps*, in "Transactions of the American Mathematical Society", February 2016, vol. 368, n^{o} 2, Also arXiv:1201.1536 [*DOI :* 10.1090/S0002-9947-2015-06413-7], https://hal.inria.fr/hal-00783682

[4] M. AKIAN, S. GAUBERT, C. WALSH.*The max-plus Martin boundary*, in "Doc. Math.", 2009, vol. 14, p. 195–240

[5] X. ALLAMIGEON, P. BENCHIMOL, S. GAUBERT, M. JOSWIG.*Combinatorial simplex algorithms can solve mean payoff games*, in "SIAM J. Opt.", 2015, vol. 24, n^{o} 4, p. 2096–2117

[6] X. ALLAMIGEON, P. BENCHIMOL, S. GAUBERT, M. JOSWIG.*Log-barrier interior point methods are not strongly polynomial*, in "SIAM Journal on Applied Algebra and Geometry", 2018, vol. 2, n^{o} 1, p. 140-178, https://arxiv.org/abs/1708.01544 - This paper supersedes arXiv:1405.4161. 31 pages, 5 figures, 1 table [*DOI :* 10.1137/17M1142132], https://hal.inria.fr/hal-01674959

[7] X. ALLAMIGEON, S. GAUBERT, E. GOUBAULT, S. PUTOT, N. STOTT.*A scalable algebraic method to infer quadratic invariants of switched systems*, in "Proceedings of the International Conference on Embedded Software (EMSOFT)", 2015, Best paper award. The extended version of this conference article appeared in ACM Trans. Embed. Comput. Syst., 15(4):69:1–69:20, September 2016

[8] J. BOLTE, S. GAUBERT, G. VIGERAL.*Definable zero-sum stochastic games*, in "Mathematics of Operations Research", 2014, vol. 40, n^{o} 1, p. 171–191, Also arXiv:1301.1967, http://dx.doi.org/10.1287/moor.2014.0666

[9] S. GAUBERT, T. LEPOUTRE.*Discrete limit and monotonicity properties of the Floquet eigenvalue in an age structured cell division cycle model*, in "J. Math. Biol.", 2015, http://dx.doi.org/10.1007/s00285-015-0874-3

[10] S. GAUBERT, G. VIGERAL.*A maximin characterization of the escape rate of nonexpansive mappings in metrically convex spaces*, in "Math. Proc. of Cambridge Phil. Soc.", 2012, vol. 152, p. 341–363, Also arXiv:1012.4765, http://dx.doi.org/10.1017/S0305004111000673

[11] C. WALSH.*The horofunction boundary and isometry group of the Hilbert geometry*, in "Handbook of Hilbert Geometry", IRMA Lectures in Mathematics and Theoretical Physics, European Mathematical Society, 2014, vol. 22, https://hal.inria.fr/hal-00782827

Publications of the year

Doctoral Dissertations and Habilitation Theses

[12] E. FODJO.*Algorithms for the resolution of stochastic control problems in high dimension by using probabilistic and max-plus methods*, Université Paris-Saclay, July 2018, https://pastel.archives-ouvertes.fr/tel-01891835

[13] M. SKOMRA.*Spectraèdres tropicaux : application à la programmation semi-définie et aux jeux à paiement moyen*, Université Paris-Saclay, December 2018, https://pastel.archives-ouvertes.fr/tel-01958741

Articles in International Peer-Reviewed Journal

[14] M. AKIAN, S. GAUBERT, A. HOCHART.*Generic uniqueness of the bias vector of finite stochastic games with perfect information*, in "Journal of Mathematical Analysis and Applications", 2018, vol. 457, p. 1038-1064, https://arxiv.org/abs/1610.09651 [*DOI :* 10.1016/J.JMAA.2017.07.017], https://hal.inria.fr/hal-01425543

[15] M. AKIAN, S. GAUBERT, A. HOCHART.*Minimax representation of nonexpansive functions and application to zero-sum recursive games*, in "Journal of Convex Analysis", February 2018, vol. 25, n° 1, https://arxiv.org/abs/1605.04518 , https://hal.inria.fr/hal-01425551

[16] M. AKIAN, S. GAUBERT, A. NIV.*Tropical compound matrix identities*, in "Linear Algebra and its Applications", August 2018, vol. 551, p. 162-206, https://arxiv.org/abs/1702.00980 [*DOI :* 10.1016/J.LAA.2018.04.006], https://hal.inria.fr/hal-01469638

[17] X. ALLAMIGEON, P. BENCHIMOL, S. GAUBERT, M. JOSWIG.*Log-barrier interior point methods are not strongly polynomial*, in "SIAM Journal on Applied Algebra and Geometry", 2018, vol. 2, n° 1, p. 140-178, https://arxiv.org/abs/1708.01544 - This paper supersedes arXiv:1405.4161. 31 pages, 5 figures, 1 table [*DOI :* 10.1137/17M1142132], https://hal.inria.fr/hal-01674959

[18] X. ALLAMIGEON, S. GAUBERT, M. SKOMRA.*Solving generic nonarchimedean semidefinite programs using stochastic game algorithms*, in "Journal of Symbolic Computation", 2018, vol. 85, p. 25-54, An abridged version of this article appeared in the proceedings of ISSAC 2016 [*DOI :* 10.1016/J.JSC.2017.07.002], https://hal.inria.fr/hal-01674494

[19] X. ALLAMIGEON, S. GAUBERT, M. SKOMRA.*The tropical analogue of the Helton–Nie conjecture is true*, in "Journal of Symbolic Computation", March 2019, vol. 91, p. 129-148, https://arxiv.org/abs/1801.02089 - The present results were announced in the conference MEGA2017 (INTERNATIONAL CONFERENCE ON EFFECTIVE METHODS IN ALGEBRAIC GEOMETRY, Université Nice Sophia Antipolis, June 2017, https://mega2017.inria.fr/) [*DOI :* 10.1016/J.JSC.2018.06.017], https://hal.inria.fr/hal-01674497

[20] X. ALLAMIGEON, R. D. KATZ.*A Formalization of Convex Polyhedra Based on the Simplex Method*, in "Journal of Automated Reasoning", August 2018, https://hal.archives-ouvertes.fr/hal-01967575

[21] D. P. CASTELLA, S. GAUBERT.*Group algebra in characteristic 1 and invariant distances over finite groups*, in "Mathematische Zeitschrift", June 2018, vol. 289, p. 695-709 [*DOI :* 10.1007/S00209-017-1971-3], https://hal.inria.fr/hal-01674503

[22] S. GAUBERT, A. NIV.*Tropical totally positive matrices*, in "Journal of Algebra", December 2018, vol. 515, p. 511-544, https://arxiv.org/abs/1606.00238 - arXiv:1606.00238 [*DOI :* 10.1016/J.JALGEBRA.2018.07.005], https://hal.inria.fr/hal-01423747

[23] P. JACQUOT, O. BEAUDE, S. GAUBERT, N. OUDJANE.*Analysis and Implementation of a Hourly Billing Mechanism for Demand Response Management*, in "IEEE Transactions on Smart Grid", July 2018, https://arxiv.org/abs/1712.08622 [*DOI :* 10.1109/TSG.2018.2855041], https://hal.inria.fr/hal-01675692

[24] N. STOTT.*Maximal lower bounds in the Löwner order*, in "Proceedings of the AMS", 2018, https://arxiv.org/abs/1612.05664 - eprint: arXiv:1612.05664 [math.RA] [*DOI :* 10.1090/PROC/13785], https://hal.inria.fr/hal-01423497

[25] C. WALSH.*Gauge-reversing maps on cones, and Hilbert and Thompson isometries*, in "Geometry and Topology", 2018, vol. 22, n^{o} 1, p. 55-104, https://arxiv.org/abs/1312.7871 - Preprint arXiv:1312.7871, https://hal.inria.fr/hal-00930929

[26] C. WALSH.*Hilbert and Thompson geometries isometric to infinite-dimensional Banach spaces*, in "Annales de l'Institut Fourier", 2018, vol. 68, n^{o} 5, p. 1831-1877, https://arxiv.org/abs/1610.07508 - See also arXiv:1610.07508, https://hal.inria.fr/hal-01249343

International Conferences with Proceedings

[27] X. ALLAMIGEON, S. GAUBERT, R. KATZ, M. SKOMRA.*Condition numbers of stochastic mean payoff games and what they say about nonarchimedean semidefinite programming*, in "23rd International Symposium on Mathematical Theory of Networks and Systems", Hong-Kong, France, July 2018, https://arxiv.org/abs/1802.07712 - 14 pages, 2 figures, https://hal.inria.fr/hal-01967555

[28] S. GAUBERT, D. JONES.*Tropical cellular automata : why urban fires propagate according to polyhedral balls*, in "Cellular Automata and Discrete Complex Systems, 24th IFIP WG 1.5 International Workshop, AUTOMATA 2018 Ghent, Belgium, June 20-22, 2018 (Exploratory Papers)", Ghent, Belgium, June 2018, https://hal.inria.fr/hal-01967561

[29] S. GAUBERT, N. STOTT.*A convergent hierarchy of non-linear eigenproblems to compute the joint spectral radius of nonnegative matrices*, in "23rd International Symposium on Mathematical Theory of Networks and Systems", Hong-Kong, France, July 2018, https://arxiv.org/abs/1805.03284 - 18 pages, https://hal.inria.fr/hal-01967552

[30] P. JACQUOT, C. WAN.*Routing Game on Parallel Networks: the Convergence of Atomic to Nonatomic*, in "CDC 2018 - IEEE 57th Conference on Decision and Control", Miami, United States, Proceedings of the 57th IEEE Conference on Decision and Control, IEEE, December 2018, vol. 1, https://hal.archives-ouvertes.fr/hal-01762547

Conferences without Proceedings

[31] X. ALLAMIGEON.*First steps in the formalization of convex polyhedra in Coq*, in "International Congress on Mathematical Software", South Bend, United States, July 2018, https://hal.archives-ouvertes.fr/hal-01967576

Scientific Books (or Scientific Book chapters)

[32] M. AKIAN, E. FODJO.*From a monotone probabilistic scheme to a probabilistic max-plus algorithm for solving Hamilton-Jacobi-Bellman equations*, in "Hamilton-Jacobi-Bellman Equations: Numerical Methods and Applications in Optimal Control", D. KALISE, K. KUNISCH, Z. RAO (editors), De Gruyter, August 2018, https://arxiv.org/abs/1709.09049 , https://hal.inria.fr/hal-01675067

[33] M. AKIAN, E. FODJO.*Probabilistic max-plus schemes for solving Hamilton-Jacobi-Bellman equations*, in "Numerical Methods for Optimal Control Problems", M. FALCONE, R. FERRETTI, L. GRUNE, W. MCENEANEY (editors), INDAM Series, Springer, February 2019, https://arxiv.org/abs/1801.01780 , https://hal.inria.fr/hal-01675068

Other Publications

[34] M. AKIAN, M. BOUHTOU, J.-B. EYTARD, S. GAUBERT.*A bilevel optimization model for load balancing in mobile networks through price incentives*, October 2018, working paper or preprint, https://hal.inria.fr/hal-01972785

[35] M. AKIAN, J.-P. CHANCELIER, B. TRAN.*A stochastic algorithm for deterministic multistage optimization problems*, December 2018, https://arxiv.org/abs/1810.12870 - working paper or preprint, https://hal.inria.fr/hal-01964189

[36] M. AKIAN, S. GAUBERT, A. HOCHART.*A game theory approach to the existence and uniqueness of nonlinear Perron-Frobenius eigenvectors*, December 2018, https://arxiv.org/abs/1812.09871 - working paper or preprint, https://hal.inria.fr/hal-01967495

[37] R. BHATIA, S. GAUBERT, T. JAIN.*Matrix versions of the Hellinger distance*, January 2019, https://arxiv.org/abs/1901.01378 - 25 pages, https://hal.inria.fr/hal-01973935

[38] G. C. CALAFIORE, S. GAUBERT, C. POSSIERI.*Log-sum-exp neural networks and posynomial models for convex and log-log-convex data*, January 2019, https://arxiv.org/abs/1806.07850 - working paper or preprint, https://hal.inria.fr/hal-01969634

[39] S. FRIEDLAND, S. GAUBERT.*Inequalities for the spectral radii and spectral norms of nonnegative tensors*, January 2019, https://arxiv.org/abs/1804.00204 - 28 pages, https://hal.inria.fr/hal-01969021

[40] S. GAUBERT, M. MACCAIG.*Approximating the Volume of Tropical Polytopes is Difficult*, January 2018, https://arxiv.org/abs/1706.06467 - working paper or preprint, https://hal.inria.fr/hal-01675715

[41] T. HAETTEL, A.-S. SCHILLING, A. WIENHARD, C. WALSH.*Horofunction Compactifications of Symmetric Spaces*, June 2018, https://arxiv.org/abs/1705.05026 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01814137

[42] P. JACQUOT, C. WAN.*Nonsmooth Aggregative Games with Coupling Constraints and Infinitely Many Classes of Players*, October 2018, https://arxiv.org/abs/1806.06230 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01904527

[43] H. LE CADRE, P. JACQUOT, C. WAN, C. ALASSEUR.*Peer-to-Peer Electricity Market Analysis: From Variational to Generalized Nash Equilibrium*, December 2018, https://arxiv.org/abs/1812.02301 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01944644

[44] C. WALSH.*Order antimorphisms of finite-dimensional cones*, January 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01673777

[45] C. WAN, P. JACQUOT, O. BEAUDE, N. OUDJANE.*Efficient Estimation of Equilibria of Large Congestion Games with Heterogeneous Players*, October 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01904546

References in notes

[46] M. AKIAN, E. FODJO.*A probabilistic max-plus numerical method for solving stochastic control problems*, in "55th Conference on Decision and Control (CDC 2016)", Las Vegas, United States, December 2016, Also arXiv:1605.02816, https://hal.inria.fr/hal-01425344

[47] M. AKIAN, S. GAUBERT.*Spectral theorem for convex monotone homogeneous maps, and ergodic control*, in "Nonlinear Anal.", 2003, vol. 52, n^{o} 2, p. 637–679, http://dx.doi.org/10.1016/S0362-546X(02)00170-0

[48] M. AKIAN, S. GAUBERT.*Policy iteration for perfect information stochastic mean payoff games with bounded first return times is strongly polynomial*, 2013, Preprint arXiv:1310.4953, 17 pages, http://hal.inria.fr/hal-00881207

[49] M. AKIAN, S. GAUBERT, A. GUTERMAN.*Linear independence over tropical semirings and beyond*, in "Proceedings of the International Conference on Tropical and Idempotent Mathematics", G. LITVINOV, S. SERGEEV (editors), Contemporary Mathematics, American Mathematical Society, 2009, vol. 495, p. 1-38, http://www.arxiv.org/abs/0812.3496

[50] M. AKIAN, S. GAUBERT, A. GUTERMAN.*Tropical polyhedra are equivalent to mean payoff games*, in "Internat. J. Algebra Comput.", 2012, vol. 22, n^{o} 1, 1250001, 43 [*DOI :* 10.1142/S0218196711006674], http://arxiv.org/abs/0912.2462

[51] M. AKIAN, S. GAUBERT, A. GUTERMAN.*Tropical Cramer Determinants Revisited*, in "Tropical and Idempotent Mathematics and Applications", G. LITVINOV, S. SERGEEV (editors), Contemporary Mathematics, AMS, 2014, vol. 616, 45, See also arXiv:1309.6298, https://hal.inria.fr/hal-00881203

[52] M. AKIAN, S. GAUBERT, M. SHARIFY.*Log-majorization of the moduli of the eigenvalues of a matrix polynomial by tropical roots*, in "Linear Algebra and its Applications", 2017, Also arXiv:1304.2967 [*DOI :* 10.1016/J.LAA.2016.11.004], https://hal.inria.fr/hal-00881196

[53] X. ALLAMIGEON, P. BENCHIMOL, S. GAUBERT.*The tropical shadow-vertex algorithm solves mean payoff games in polynomial time on average*, in "ICALP 2014", Copenhagen, France, J. ESPARZA, P. FRAIGNIAUD, T. HUSFELDT, E. KOUTSOUPIAS (editors), 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part I, Springer, July 2014, vol. 8572, 12 [*DOI :* 10.1007/978-3-662-43948-7_8], https://hal.inria.fr/hal-01096447

[54] X. ALLAMIGEON, P. BENCHIMOL, S. GAUBERT, M. JOSWIG.*Long and winding central paths*, May 2014, Preprint arXiv:1405.4161, v2 May 2015, https://hal.inria.fr/hal-01096452

[55] X. ALLAMIGEON, V. BOEUF, S. GAUBERT.*Performance evaluation of an emergency call center: tropical polynomial systems applied to timed Petri nets*, in "13th International Conference, Formal Modeling and Analysis of Timed Systems (FORMATS 2015)", Madrid, Spain, Formal Modeling and Analysis of Timed Systems, Springer, September 2015, vol. 9268 [*DOI :* 10.1007/978-3-319-22975-1_2], https://hal.inria.fr/hal-01248814

[56] X. ALLAMIGEON, V. BOEUF, S. GAUBERT.*Stationary solutions of discrete and continuous Petri nets with priorities*, in "Performance Evaluation", August 2017, vol. 113, p. 1 - 12, https://arxiv.org/abs/1612.07661 [*DOI :* 10.1016/J.PEVA.2017.04.007], https://hal.inria.fr/hal-01674492

[57] X. ALLAMIGEON, S. GAUBERT, E. GOUBAULT.*Inferring Min and Max Invariants Using Max-plus Polyhedra*, in "Proceedings of the 15th International Static Analysis Symposium (SAS'08)", Valencia, Spain, LNCS, Springer, 2008, vol. 5079, p. 189–204, http://dx.doi.org/10.1007/978-3-540-69166-2_13

[58] X. ALLAMIGEON, S. GAUBERT, E. GOUBAULT.*Computing the Vertices of Tropical Polyhedra using Directed Hypergraphs*, in "Discrete Comp. Geom.", 2012, Published on line [*DOI :* 10.1007/s00454-012-9469-6], http://fr.arxiv.org/abs/0904.3436v3

[59] X. ALLAMIGEON, S. GAUBERT, M. SKOMRA.*Solving Generic Nonarchimedean Semidefinite Programs Using Stochastic Game Algorithms*, in "ISSAC '16: International Symposium on Symbolic and Algebraic Computation", Waterloo, France, Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation (ISSAC'16), ACM, July 2016, Also arXiv:1603.06916 [*DOI :* 10.1145/2930889.2930935], https://hal.inria.fr/hal-01422638

[60] X. ALLAMIGEON, S. GAUBERT, M. SKOMRA.*Tropical spectrahedra*, October 2016, arXiv:1610.06746, https://hal.inria.fr/hal-01422639

[61] P. ANDY, W. FAISAL, F. BONNANS.*MIDAS: A Mixed Integer Dynamic Approximation Scheme*, Inria, 2016, https://hal.inria.fr/hal-01401950

[62] F. BACCELLI, G. COHEN, G.-J. OLSDER, J.-P. QUADRAT.*Synchronization and linearity: an algebra for discrete event systems*, Wiley, 1992

[63] G. BARLES, S. MIRRAHIMI, B. PERTHAME.*Concentration in Lotka-Volterra parabolic or integral equations: a general convergence result*, in "Methods Appl. Anal.", 2009, vol. 16, n^o 3, p. 321–340, http://dx.doi.org/10.4310/MAA.2009.v16.n3.a4

[64] R. BHATIA, S. GAUBERT, T. JAIN.*Matrix versions of the Hellinger distance*, 2019, To appear in Letters in Math. Physics

[65] V. BOEUF, P. ROBERT.*A Stochastic Analysis of a Network with Two Levels of Service*, August 2017, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01583704

[66] F. BONNANS, S. GAUBERT.*Recherche opérationnelle. Aspects mathématiques et applications*, Ellipse, March 2016, 391, https://hal.inria.fr/hal-01422645

[67] P. BUTKOVIČ.*Max-algebra: the linear algebra of combinatorics?*, in "Linear Algebra and its applications", 2003, vol. 367, p. 313–335

[68] P. BUTKOVIČ.*Max-linear systems: theory and algorithms*, Springer Monographs in Mathematics, Springer-Verlag London, Ltd., London, 2010, xviii+272, http://dx.doi.org/10.1007/978-1-84996-299-5

[69] J. COCHET-TERRASSON, G. COHEN, S. GAUBERT, M. MC GETTRICK, J.-P. QUADRAT.*Numerical computation of spectral elements in max-plus algebra*, in "Proc. of the IFAC Conference on System Structure and Control", Nantes, July 1998

[70] G. COHEN, S. GAUBERT, J.-P. QUADRAT.*Max-plus algebra and system theory: where we are and where to go now*, in "Annual Reviews in Control", 1999, vol. 23, p. 207–219

[71] A. CONNES.*Trace formula in noncommutative geometry and the zeros of the Riemann zeta function*, in "Selecta Math. (N.S.)", 1999, vol. 5, n^o 1, p. 29–106

[72] A. CONNES, C. CONSANI.*The Arithmetic Site*, in "Comptes Rendus Mathématiques", 2014, vol. Ser. I 352, p. 971–975

[73] A. CONNES, C. CONSANI.*Geometry of the arithmetic site*, 2015, arXiv:1502.05580

[74] A. CONNES, C. CONSANI.*Geometry of the arithmetic site*, in "Adv. Math.", 2016, vol. 291, p. 274–329

[75] P. COUSOT, R. COUSOT.*Abstract Interpretation: A unified lattice model for static analysis of programs by construction of approximations of fixed points*, in "Principles of Programming Languages 4", 1977, p. 238–252

[76] J. DE LOERA, B. STURMFELS, C. VINZANT.*The central curve in linear programming*, in "Foundations of Computational Mathematics", 2012, vol. 12, n° 4, p. 509–540

[77] J.-P. DEDIEU, G. MALAJOVICH, M. SHUB.*On the Curvature of the Central Path of Linear Programming Theory*, in "Foundations of Computational Mathematics", 2005, vol. 5, n° 2, p. 145–171

[78] A. DEZA, T. TERLAKY, Y. ZINCHENKO.*Polytopes and arrangements: diameter and curvature*, in "Operations Research Letters", 2008, vol. 36, n° 2, p. 215–222

[79] A. DEZA, T. TERLAKY, Y. ZINCHENKO.*Central path curvature and iteration-complexity for redundant Klee-Minty cubes*, in "Advances in applied mathematics and global optimization", New York, Adv. Mech. Math., Springer, 2009, vol. 17, p. 223–256, http://dx.doi.org/10.1007/978-0-387-75714-8_7

[80] J. B. EYTARD, M. AKIAN, M. BOUHTOU, S. GAUBERT.*A bilevel optimization model for load balancing in mobile networks through price incentives*, in "WiOpt 2017 - 15th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks", Paris, France, IEEE, May 2017, p. 1-8 [*DOI :* 10.23919/WIOPT.2017.7959902], https://hal.inria.fr/hal-01649042

[81] J.-B. EYTARD.*A tropical geometry and discrete convexity approach to bilevel programming : application to smart data pricing in mobile telecommunication networks*, Université Paris-Saclay, November 2018, https://pastel.archives-ouvertes.fr/tel-01972391

[82] A. FAHIM, N. TOUZI, X. WARIN.*A probabilistic numerical method for fully nonlinear parabolic PDEs*, in "Ann. Appl. Probab.", 2011, vol. 21, n° 4, p. 1322–1364, http://dx.doi.org/10.1214/10-AAP723

[83] M. FARBER.*Invitation to Topological Robotics*, Zurich lectures in advanced mathematics, European Mathematical Society, 2008

[84] M. FARBER, E. GOUBAULT, A. SAGNIER.*Directed topological complexity*, 2018, https://hal.archives-ouvertes.fr/hal-01970569

[85] A. FATHI, A. SICONOLFI.*Existence of C^1 critical subsolutions of the Hamilton-Jacobi equation*, in "Invent. Math.", 2004, vol. 155, n° 2, p. 363–388, http://dx.doi.org/10.1007/s00222-003-0323-6

[86] O. FERCOQ, M. AKIAN, M. BOUHTOU, S. GAUBERT.*Ergodic control and polyhedral approaches to PageRank optimization*, in "IEEE Trans. Automat. Control", 2013, vol. 58, n° 1, p. 134–148, http://dx.doi.org/10.1109/TAC.2012.2226103

[87] W. FLEMING, W. MCENEANEY.*A max-plus based algorithm for an HJB equation of non-linear filtering*, in "SIAM J. Control and Opt.", 2000, p. 683–710

[88] S. FOMIN, A. ZELEVINSKY.*Cluster algebras. I. Foundations*, in "J. Amer. Math. Soc.", 2002, vol. 15, n^o 2, p. 497–529, http://arxiv.org/abs/math.RT/0104151

[89] S. FRIEDLAND, S. GAUBERT, L. HAN.*Perron–Frobenius theorem for nonnegative multilinear forms and extensions*, in "Linear Algebra Appl.", 2013, vol. 438, n^o 2, p. 738–749 [*DOI :* 10.1016/J.LAA.2011.02.042], http://hal.inria.fr/hal-00782755

[90] S. GAUBERT, E. GOUBAULT, A. TALY, S. ZENNOU.*Static Analysis by Policy Iteration in Relational Domains*, in "Proceedings of the Proc. of the 16th European Symposium on Programming (ESOP'07)", Braga (Portugal), LNCS, Springer, 2007, vol. 4421, p. 237–252, http://dx.doi.org/10.1007/978-3-540-71316-6_17

[91] S. GAUBERT, J. GUNAWARDENA.*The Perron-Frobenius Theorem for Homogeneous, Monotone Functions*, in "Trans. of AMS", 2004, vol. 356, n^o 12, p. 4931-4950, http://www.ams.org/tran/2004-356-12/S0002-9947-04-03470-1/home.html

[92] S. GAUBERT, W. MCENEANEY, Z. QU.*Curse of dimensionality reduction in max-plus based approximation methods: theoretical estimates and improved pruning algorithms*, in "Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC 11)", Orlando, FL, USA, December 2011, p. 1054-1061, http://arxiv.org/abs/1109.5241

[93] S. GAUBERT, M. SHARIFY.*Tropical scaling of polynomial matrices*, in "Positive systems", Berlin, Lecture Notes in Control and Inform. Sci., Springer, 2009, vol. 389, p. 291–303, http://dx.doi.org/10.1007/978-3-642-02894-6_28

[94] S. GAUBERT, N. STOTT.*Tropical Kraus maps for optimal control of switched systems*, in "Proceedings of the 57th IEEE Annual Conference on Decision and Control (CDC)", Melbourne, Australia, 2017, https://arxiv.org/abs/1706.04471 - 15 pages, https://hal.inria.fr/hal-01674496

[95] T. M. GAWLITZA, H. SEIDL, A. ADJÉ, S. GAUBERT, E. GOUBAULT.*Abstract interpretation meets convex optimization*, in "J. Symbolic Comput.", 2012, vol. 47, n^o 12, p. 1416–1446, Special issue on Invariant generation and reasoning about loops, http://dx.doi.org/10.1016/j.jsc.2011.12.048

[96] I. M. GELFAND, M. M. KAPRANOV, A. V. ZELEVINSKY.*Discriminants, resultants and multidimensional determinants*, Modern Birkhäuser Classics, Birkhäuser Boston, Inc., Boston, MA, 2008, x+523, Reprint of the 1994 edition

[97] M. GRANDIS.*Directed Algebraic Topology, Models of non-reversible worlds*, Cambridge University Press, 2009

[98] S. HAMMARLING, C. J. MUNRO, F. TISSEUR.*An algorithm for the complete solution of quadratic eigenvalue problems*, in "ACM Trans. Math. Software", 2013, vol. 39, n^o 3, Art. 18, 19, http://dx.doi.org/10.1145/2450153.2450156

[99] B. HEIDERGOTT, G. J. OLSDER, J. VAN DER WOUDE.*Max Plus at Work: Modeling and Analysis of Synchronized Systems: A Course on Max-Plus Algebra and Its Applications*, Princeton, 2005

[100] H. ISHII, H. MITAKE.*Representation formulas for solutions of Hamilton-Jacobi equations with convex Hamiltonians*, in "Indiana Univ. Math. J.", 2007, vol. 56, n$^{\text{o}}$ 5, p. 2159–2183, http://dx.doi.org/10.1512/iumj.2007.56.3048

[101] I. ITENBERG, G. MIKHALKIN, E. SHUSTIN.*Tropical algebraic geometry*, Oberwolfach Seminars, Birkhäuser Verlag, Basel, 2007, vol. 35, viii+103

[102] P. JACQUOT, O. BEAUDE, S. GAUBERT, N. OUDJANE.*Demand Side Management in the Smart Grid: an Efficiency and Fairness Tradeoff*, in "7th IEEE International Conference on Innovative Smart Grid Technologies", Torino, France, August 2017, https://arxiv.org/abs/1711.11129 , https://hal.inria.fr/hal-01675658

[103] H. KAISE, W. M. MCENEANEY.*Idempotent expansions for continuous-time stochastic control: compact control space*, in "Proceedings of the 49th IEEE Conference on Decision and Control", Atlanta, Dec. 2010

[104] V. KOLOKOLTSOV, V. MASLOV.*Idempotent analysis and applications*, Kluwer Acad. Publisher, 1997

[105] B. LEMMENS, R. NUSSBAUM.*Nonlinear Perron-Frobenius theory*, Cambridge Tracts in Mathematics, Cambridge University Press, Cambridge, 2012, vol. 189, xii+323, http://dx.doi.org/10.1017/CBO9781139026079

[106] Q. LU, M. MADSEN, M. MILATA, S. RAVN, U. FAHRENBERG, K. G. LARSEN.*Reachability Analysis for Timed Automata using Max-Plus Algebra*, in "J. Logic Alg. Prog.", 2012, vol. 81, n$^{\text{o}}$ 3, p. 298-313

[107] V. MASLOV.*Méthodes Operatorielles*, Edition Mir, Moscou, 1987

[108] W. MCENEANEY, A. DESHPANDE, S. GAUBERT.*Curse-of-Complexity Attenuation in the Curse-of-Dimensionality-Free Method for HJB PDEs*, in "Proc. of the 2008 American Control Conference", Seattle, Washington, USA, June 2008

[109] W. M. MCENEANEY, H. KAISE, S. H. HAN.*Idempotent Method for Continuous-time Stochastic Control and Complexity Attenuation*, in "Proceedings of the 18th IFAC World Congress, 2011", Milano, Italie, 2011, p. 3216-3221

[110] W. M. MCENEANEY.*Max-plus methods for nonlinear control and estimation*, Systems & Control: Foundations & Applications, Birkhäuser Boston Inc., Boston, MA, 2006, xiv+241

[111] W. M. MCENEANEY.*A curse-of-dimensionality-free numerical method for solution of certain HJB PDEs*, in "SIAM J. Control Optim.", 2007, vol. 46, n$^{\text{o}}$ 4, p. 1239–1276, http://dx.doi.org/10.1137/040610830

[112] J.-F. MERTENS, S. SORIN, S. ZAMIR.*Repeated Games*, Cambridge, 2015

[113] G. MIKHALKIN.*Enumerative tropical algebraic geometry in $\mathbb{R}^2$*, in "J. Amer. Math. Soc.", 2005, vol. 18, n$^{\text{o}}$ 2, p. 313–377, http://dx.doi.org/10.1090/S0894-0347-05-00477-7

[114] R. H. MÖHRING, M. SKUTELLA, F. STORK.*Scheduling with AND/OR precedence constraints*, in "SIAM J. Comput.", 2004, vol. 33, n$^{\text{o}}$ 2, p. 393–415, http://dx.doi.org/10.1137/S009753970037727X

[115] A. PAPADOPOULOS.*Metric spaces, convexity and non-positive curvature*, IRMA Lectures in Mathematics and Theoretical Physics, Second, European Mathematical Society (EMS), Zürich, 2014, vol. 6, xii+309, http://dx.doi.org/10.4171/132

[116] M. V. F. PEREIRA, L. M. V. G. PINTO.*Multi-stage stochastic optimization applied to energy planning*, in "Math. Programming", 1991, vol. 52, n^{o} 2, Ser. B, p. 359–375, http://dx.doi.org/10.1007/BF01582895

[117] J.-E. PIN.*Tropical Semirings*, in "Idempotency", J. GUNAWARDENA (editor), Publications of the Isaac Newton Institute, Cambridge University Press, 1998, vol. 11, p. 50–69

[118] Z. QU.*Théorie de Perron-Frobenius non linéaire et méthodes numériques max-plus pour la résolution d'équations d'Hamilton-Jacobi*, Ecole Polytechnique X, October 2013, http://hal.inria.fr/pastel-00927122

[119] D. REEB, M. J. KASTORYANO, M. M. WOLF.*Hilbert's projective metric in quantum information theory*, in "J. Math. Phys.", 2011, vol. 52, n^{o} 8, 082201, 33, http://dx.doi.org/10.1063/1.3615729

[120] J. RICHTER-GEBERT, B. STURMFELS, T. THEOBALD.*First steps in tropical geometry*, in "Idempotent mathematics and mathematical physics", Providence, RI, Contemp. Math., Amer. Math. Soc., 2005, vol. 377, p. 289–317

[121] A. SAGNIER.*An arithmetic site of Connes-Consani type for imaginary quadratic fields with class number 1*, 2018, submitted

[122] G. SAGNOL, S. GAUBERT, M. BOUHTOU.*Optimal monitoring on large networks by Successive c-Optimal Designs*, in "Proceedings of the 22nd international teletraffic congress (ITC22), Amsterdam, The Netherlands, September", IEEE, 2010, http://dx.doi.org/10.1109/ITC.2010.5608717

[123] S. SANKARANARAYANAN, H. SIPMA, Z. MANNA.*Scalable Analysis of Linear Systems using Mathematical Programming*, in "VMCAI", LNCS, 2005, vol. 3385

[124] C. SCHEIDERER.*Semidefinitely representable convex sets*, 2016, arXiv:1612.07048v2

[125] R. SEPULCHRE, A. SARLETTE, P. ROUCHON.*Consensus in noncommutative spaces*, in "Proceedings of the 49th IEEE Conference on Decision and Control", Atlanta, USA, Dec 2010, p. 6596-6601, http://dx.doi.org/10.1109/CDC.2010.5717072

[126] A. SIDFORD, M. WANG, X. WU, Y. YE.*Variance reduced value iteration and faster algorithms for solving Markov decision processes*, in "Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms", SIAM, Philadelphia, PA, 2018, p. 770–787, https://doi.org/10.1137/1.9781611975031.50

[127] I. SIMON.*Limited subsets of a free monoid*, in "Proc. 19th Annual Symposium on Foundations of Computer Science", Piscataway, NJ, 1978, p. 143–150

[128] S. SMALE.*Mathematical problems for the next century*, in "Math. Intelligencer", 1998, vol. 20, n^{o} 2, p. 7–15, http://dx.doi.org/10.1007/BF03025291

[129] H. L. SMITH.*Monotone dynamical systems*, Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 1995, vol. 41, x+174, An introduction to the theory of competitive and cooperative systems

[130] N. M. TRAN, J. YU.*Product-Mix Auctions and Tropical Geometry*, 2015

[131] O. VIRO.*Dequantization of real algebraic geometry on logarithmic paper*, in "European Congress of Mathematics, Vol. I (Barcelona, 2000)", Basel, Progr. Math., Birkhäuser, 2001, vol. 201, p. 135–146

[132] O. VIRO.*Dequantization of real algebraic geometry on logarithmic paper*, in "European Congress of Mathematics, Vol. I (Barcelona, 2000)", Basel, Progr. Math., Birkhäuser, 2001, vol. 201, p. 135–146, http://arxiv.org/abs/math.AG/0005163

[133] C. WALSH.*Gauge-reversing maps on cones, and Hilbert and Thompson isometries*, 2013, 36 pages, 3 figures, http://hal.inria.fr/hal-00930929

Project-Team XPOP

Statistical modelling for life sciences

IN COLLABORATION WITH: Centre de Mathématiques Appliquées (CMAP)

IN PARTNERSHIP WITH:
Ecole Polytechnique

RESEARCH CENTER
Saclay - Île-de-France

THEME
Modeling and Control for Life Sciences

Table of contents

Project-Team XPOP

Creation of the Team: 2016 January 01, updated into Project-Team: 2017 July 01

Keywords:

Computer Science and Digital Science:

A3.1.1. - Modeling, representation
A3.2.3. - Inference
A3.3. - Data and knowledge analysis
A3.3.1. - On-line analytical processing
A3.3.2. - Data mining
A3.3.3. - Big data analysis
A3.4.1. - Supervised learning
A3.4.2. - Unsupervised learning
A3.4.4. - Optimization and learning
A3.4.5. - Bayesian methods
A3.4.6. - Neural networks
A3.4.7. - Kernel methods
A3.4.8. - Deep learning
A5.9.2. - Estimation, modeling
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.2.2. - Numerical probability
A6.2.3. - Probabilistic methods
A6.2.4. - Statistical methods
A6.3.3. - Data processing
A6.3.5. - Uncertainty Quantification

Other Research Topics and Application Domains:

B1.1.4. - Genetics and genomics
B1.1.7. - Bioinformatics
B1.1.10. - Systems and synthetic biology
B2.2.3. - Cancer
B2.2.4. - Infectious diseases, Virology
B2.4.1. - Pharmaco kinetics and dynamics
B9.1.1. - E-learning, MOOC

1. Team, Visitors, External Collaborators

Research Scientists

Marc Lavielle [Team leader, Inria, Senior Researcher]
Laetitia Le [APHP - Ecole polytechnique, Researcher, from Nov 2018]

Faculty Members

Julie Josse [Ecole polytechnique, Associate Professor]
Erwan Le Pennec [Ecole Polytechnique, Associate Professor, HDR]
Eric Moulines [Ecole polytechnique, Professor]

Technical Staff
Yao Xu [Inria]

PhD Students
Nicolas Brosse [Ecole polytechnique]
Wei Jiang [Ecole polytechnique]
Mohammed Karimi [Inria]
Genevieve Robin [Ecole polytechnique]
Marine Zulian [Dassault Systemes]

Post-Doctoral Fellow
Tom Rohmer [Inria, from Sep 2018]

Visiting Scientist
Ricardo Rios [Universidad Central de Venezuela, Sep 2018]

Administrative Assistants
Hanadi Dib [Inria, until Oct 2018]
Ines Dumontier [Inria, from Nov 2018]
Hélèna Kutniak [Inria, from Oct 2018]

2. Overall Objectives

2.1. Developing sound, useful and usable methods

The main objective of XPOP is to develop new sound and rigorous methods for statistical modeling in the field of biology and life sciences. These methods for modeling include statistical methods of estimation, model diagnostics, model building and model selection as well as methods for numerical models (systems of ordinary and partial differential equations). Historically, the key area where these methods have been used is population pharmacokinetics. However, the framework is currently being extended to sophisticated numerical models in the contexts of viral dynamics, glucose-insulin processes, tumor growth, precision medicine, spectrometry, intracellular processes, etc.

Furthermore, an important aim of XPOP is to transfer the methods developed into software packages so that they can be used in everyday practice.

2.2. Combining numerical, statistical and stochastic components of a model

Mathematical models that characterize complex biological phenomena are defined by systems of ordinary differential equations when dealing with dynamical systems that evolve with respect to time, or by partial differential equations when there is a spatial component in the model. Also, it is sometimes useful to integrate a stochastic aspect into the dynamical system in order to model stochastic intra-individual variability.

In order to use such methods, we must deal with complex numerical difficulties, generally related to resolving the systems of differential equations. Furthermore, to be able to check the quality of a model (i.e. its descriptive and predictive performances), we require data. The statistical aspect of the model is thus critical in how it takes into account different sources of variability and uncertainty, especially when data come from several individuals and we are interested in characterizing the inter-subject variability. Here, the tools of reference are mixed-effects models.

Confronted with such complex modeling problems, one of the goals of XPOP is to show the importance of combining numerical, statistical and stochastic approaches.

2.3. Developing future standards

Linear mixed-effects models have been well-used in statistics for a long time. They are a classical approach, essentially relying on matrix calculations in Gaussian models. Whereas a solid theoretical base has been developed for such models, *nonlinear* mixed-effects models (NLMEM) have received much less attention in the statistics community, even though they have been applied to many domains of interest. It has thus been the users of these models, such as pharmacometricians, who have taken them and developed methods, without really looking to develop a clean theoretical framework or understand the mathematical properties of the methods. This is why a standard estimation method in NLMEM is to linearize the model, and few people have been interested in understanding the properties of estimators obtained in this way.

Statisticians and pharmacometricians frequently realize the need to create bridges between these two communities. We are entirely convinced that this requires the development of new standards for population modeling that can be widely accepted by these various communities. These standards include the language used for encoding a model, the approach for representing a model and the methods for using it:

- **The approach.** Our approach consists in seeing a model as hierarchical, represented by a joint probability distribution. This joint distribution can be decomposed into a product of conditional distributions, each associated with a submodel (model for observations, individual parameters, etc.). Tasks required of the modeler are thus related to these probability distributions.
- **The methods.** Many tests have shown that algorithms implemented in MONOLIX are the most reliable, all the while being extremely fast. In fact, these algorithms are precisely described and published in well known statistical journals. In particular, the SAEM algorithm, used for calculating the maximum likelihood estimation of population parameters, has shown its worth in numerous situations. Its mathematical convergence has also been proven under quite general hypotheses.
- **The language.** Mlxtran is used by MONOLIX and other modeling tools and is today by far the most advanced language for representing models. Initially developed for representing pharmacometric models, its syntax also allows it to easily code dynamical systems defined by a system of ODEs, and statistical models involving continuous, discrete and survival variables. This flexibility is a true advantage both for numerical modelers and statisticians.

3. Research Program

3.1. Scientific positioning

"Interfaces" is the defining characteristic of XPOP:

The interface between statistics, probability and numerical methods. Mathematical modelling of complex biological phenomena require to combine numerical, stochastic and statistical approaches. The CMAP is therefore the right place to be for positioning the team at the interface between several mathematical disciplines.

The interface between mathematics and the life sciences. The goal of XPOP is to bring the right answers to the right questions. These answers are mathematical tools (statistics, numerical methods, etc.), whereas the questions come from the life sciences (pharmacology, medicine, biology, etc.). This is why the point of XPOP is not to take part in mathematical projects only, but also pluridisciplinary ones.

The interface between mathematics and software development. The development of new methods is the main activity of XPOP. However, new methods are only useful if they end up being implemented in a software tool. On one hand, a strong partnership with Lixoft (the spin-off company who continue developing MONOLIX) allows us to maintaining this positioning. On the other hand, several members of the team are very active in the R community and develop widely used packages.

3.2. The mixed-effects models

Mixed-effects models are statistical models with both fixed effects and random effects. They are well-adapted to situations where repeated measurements are made on the same individual/statistical unit.

Consider first a single subject i of the population. Let $y_i = (y_{ij}, 1 \leq j \leq n_i)$ be the vector of observations for this subject. The model that describes the observations y_i is assumed to be a parametric probabilistic model: let $p_Y(y_i; \psi_i)$ be the probability distribution of y_i, where ψ_i is a vector of parameters.

In a population framework, the vector of parameters ψ_i is assumed to be drawn from a population distribution $p_\Psi(\psi_i; \theta)$ where θ is a vector of population parameters.

Then, the probabilistic model is the joint probability distribution

$$p(y_i, \psi_i; \theta) = p_Y(y_i|\psi_i) p_\Psi(\psi_i; \theta) \tag{17}$$

To define a model thus consists in defining precisely these two terms.

In most applications, the observed data y_i are continuous longitudinal data. We then assume the following representation for y_i:

$$y_{ij} = f(t_{ij}, \psi_i) + g(t_{ij}, \psi_i)\varepsilon_{ij} \quad , \ 1 \leq i \leq N \quad , \quad 1 \leq j \leq n_i. \tag{18}$$

Here, y_{ij} is the observation obtained from subject i at time t_{ij}. The residual errors (ε_{ij}) are assumed to be standardized random variables (mean zero and variance 1). The residual error model is represented by function g in model (2).

Function f is usually the solution to a system of ordinary differential equations (pharmacokinetic/pharmacodynamic models, etc.) or a system of partial differential equations (tumor growth, respiratory system, etc.). This component is a fundamental component of the model since it defines the prediction of the observed kinetics for a given set of parameters.

The vector of individual parameters ψ_i is usually function of a vector of population parameters ψ_{pop}, a vector of random effects $\eta_i \sim \mathcal{N}(0, \Omega)$, a vector of individual covariates c_i (weight, age, gender, ...) and some fixed effects β.

The joint model of y and ψ depends then on a vector of parameters $\theta = (\psi_{\text{pop}}, \beta, \Omega)$.

3.3. Computational Statistical Methods

Central to modern statistics is the use of probabilistic models. To relate these models to data requires the ability to calculate the probability of the observed data: the likelihood function, which is central to most statistical methods and provides a principled framework to handle uncertainty.

The emergence of computational statistics as a collection of powerful and general methodologies for carrying out likelihood-based inference made complex models with non-standard data accessible to likelihood, including hierarchical models, models with intricate latent structure, and missing data.

In particular, algorithms previously developed by POPIX for mixed effects models, and today implemented in several software tools (especially MONOLIX) are part of these methods:

- the adaptive Metropolis-Hastings algorithm allows one to sample from the conditional distribution of the individual parameters $p(\psi_i|y_i; c_i, \theta)$,
- the SAEM algorithm is used to maximize the observed likelihood $\mathcal{L}(\theta; y) = p(y; \theta)$,
- Importance Sampling Monte Carlo simulations provide an accurate estimation of the observed log-likelihood $\log(\mathcal{L}(\theta; y))$.

Computational statistics is an area which remains extremely active today. Recently, one can notice that the incentive for further improvements and innovation comes mainly from three broad directions: the high dimensional challenge, the quest for adaptive procedures that can eliminate the cumbersome process of tuning "by hand" the settings of the algorithms and the need for flexible theoretical support, arguably required by all recent developments as well as many of the traditional MCMC algorithms that are widely used in practice.

Working in these three directions is a clear objective for XPOP.

3.4. Markov Chain Monte Carlo algorithms

While these Monte Carlo algorithms have turned into standard tools over the past decade, they still face difficulties in handling less regular problems such as those involved in deriving inference for high-dimensional models. One of the main problems encountered when using MCMC in this challenging settings is that it is difficult to design a Markov chain that efficiently samples the state space of interest.

The Metropolis-adjusted Langevin algorithm (MALA) is a Markov chain Monte Carlo (MCMC) method for obtaining random samples from a probability distribution for which direct sampling is difficult. As the name suggests, MALA uses a combination of two mechanisms to generate the states of a random walk that has the target probability distribution as an invariant measure:

1. new states are proposed using Langevin dynamics, which use evaluations of the gradient of the target probability density function;
2. these proposals are accepted or rejected using the Metropolis-Hastings algorithm, which uses evaluations of the target probability density (but not its gradient).

Informally, the Langevin dynamics drives the random walk towards regions of high probability in the manner of a gradient flow, while the Metropolis-Hastings accept/reject mechanism improves the mixing and convergence properties of this random walk.

Several extensions of MALA have been proposed recently by several authors, including fMALA (fast MALA), AMALA (anisotropic MALA), MMALA (manifold MALA), position-dependent MALA (PMALA), ...

MALA and these extensions have demonstrated to represent very efficient alternative for sampling from high dimensional distributions. We therefore need to adapt these methods to general mixed effects models.

3.5. Parameter estimation

The Stochastic Approximation Expectation Maximization (SAEM) algorithm has shown to be extremely efficient for maximum likelihood estimation in incomplete data models, and particularly in mixed effects models for estimating the population parameters. However, there are several practical situations for which extensions of SAEM are still needed:

High dimensional model: a complex physiological model may have a large number of parameters (in the order of 100). Then several problems arise:

- when most of these parameters are associated with random effects, the MCMC algorithm should be able to sample, for each of the N individuals, parameters from a high dimensional distribution. Efficient MCMC methods for high dimensions are then required.
- Practical identifiability of the model is not ensured with a limited amount of data. In other words, we cannot expect to be able to properly estimate all the parameters of the model, including the fixed effects and the variance-covariance matrix of the random effects. Then, some random effects should be removed, assuming that some parameters do not vary in the population. It may also be necessary to fix the value of some parameters (using values from the literature for instance). The strategy to decide which parameters should be fixed and which random effects should be removed remains totally empirical. XPOP aims to develop a procedure that will help the modeller to take such decisions.

Large number of covariates: the covariate model aims to explain part of the inter-patient variability of some parameters. Classical methods for covariate model building are based on comparisons with respect to some criteria, usually derived from the likelihood (AIC, BIC), or some statistical test (Wald test, LRT, etc.). In other words, the modelling procedure requires two steps: first, all possible models are fitted using some estimation procedure (e.g. the SAEM algorithm) and the likelihood of each model is computed using a numerical integration procedure (e.g. Monte Carlo Importance Sampling); then, a model selection procedure chooses the "best" covariate model. Such a strategy is only possible with a reduced number of covariates, i.e., with a "small" number of models to fit and compare.

As an alternative, we are thinking about a Bayesian approach which consists of estimating simultaneously the covariate model and the parameters of the model in a single run. An (informative or uninformative) prior is defined for each model by defining a prior probability for each covariate to be included in the model. In other words, we extend the probabilistic model by introducing binary variables that indicate the presence or absence of each covariate in the model. Then, the model selection procedure consists of estimating and maximizing the conditional distribution of this sequence of binary variables. Furthermore, a probability can be associated to any of the possible covariate models.

This conditional distribution can be estimated using an MCMC procedure combined with the SAEM algorithm for estimating the population parameters of the model. In practice, such an approach can only deal with a limited number of covariates since the dimension of the probability space to explore increases exponentially with the number of covariates. Consequently, we would like to have methods able to find a small number of variables (from a large starting set) that influence certain parameters in populations of individuals. That means that, instead of estimating the conditional distribution of all the covariate models as described above, the algorithm should focus on the most likely ones.

Fixed parameters: it is quite frequent that some individual parameters of the model have no random component and are purely fixed effects. Then, the model may not belong to the exponential family anymore and the original version of SAEM cannot be used as it is. Several extensions exist:

- introduce random effects with decreasing variances for these parameters,
- introduce a prior distribution for these fixed effects,
- apply the stochastic approximation directly on the sequence of estimated parameters, instead of the sufficient statistics of the model.

None of these methods always work correctly. Furthermore, what are the pros and cons of these methods is not clear at all. Then, developing a robust methodology for such model is necessary.

Convergence toward the global maximum of the likelihood: convergence of SAEM can strongly depend on thie initial guess when the observed likelihood has several local maxima. A kind of simulated annealing version of SAEM was previously developed and implemented in MONOLIX. The method works quite well in most situations but there is no theoretical justification and choosing the settings of this algorithm (i.e. how the temperature decreases during the iterations) remains empirical. A precise analysis of the algorithm could be very useful to better understand why it "works" in practice and how to optimize it.

Convergence diagnostic: Convergence of SAEM was theoretically demonstrated under very general hypothesis. Such result is important but of little interest in practice at the time to use SAEM in a finite amount of time, i.e. in a finite number of iterations. Some qualitative and quantitative criteria should be defined in order to both optimize the settings of the algorithm, detect a poor convergence of SAEM and evaluate the quality of the results in order to avoid using them unwisely.

3.6. Model building

Defining an optimal strategy for model building is far from easy because a model is the assembled product of numerous components that need to been evaluated and perhaps improved: the structural model, residual error model, covariate model, covariance model, etc.

How to proceed so as to obtain the best possible combination of these components? There is no magic recipe but an effort will be made to provide some qualitative and quantitative criteria in order to help the modeller for building his model.

The strategy to take will mainly depend on the time we can dedicate to building the model and the time required for running it. For relatively simple models for which parameter estimation is fast, it is possible to fit many models and compare them. This can also be done if we have powerful computing facilities available (e.g., a cluster) allowing large numbers of simultaneous runs.

However, if we are working on a standard laptop or desktop computer, model building is a sequential process in which a new model is tested at each step. If the model is complex and requires significant computation time (e.g., when involving systems of ODEs), we are constrained to limit the number of models we can test in a reasonable time period. In this context, it also becomes important to carefully choose the tasks to run at each step.

3.7. Model evaluation

Diagnostic tools are recognized as an essential method for model assessment in the process of model building. Indeed, the modeler needs to confront "his" model with the experimental data before concluding that this model is able to reproduce the data and before using it for any purpose, such as prediction or simulation for instance.

The objective of a diagnostic tool is twofold: first we want to check if the assumptions made on the model are valid or not ; then, if some assumptions are rejected, we want to get some guidance on how to improve the model.

As is the usual case in statistics, it is not because this "final" model has not been rejected that it is necessarily the "true" one. All that we can say is that the experimental data does not allow us to reject it. It is merely one of perhaps many models that cannot be rejected.

Model diagnostic tools are for the most part graphical, i.e., visual; we "see" when something is not right between a chosen model and the data it is hypothesized to describe. These diagnostic plots are usually based on the empirical Bayes estimates (EBEs) of the individual parameters and EBEs of the random effects: scatterplots of individual parameters versus covariates to detect some possible relationship, scatterplots of pairs of random effects to detect some possible correlation between random effects, plot of the empirical distribution of the random effects (boxplot, histogram,...) to check if they are normally distributed, ...

The use of EBEs for diagnostic plots and statistical tests is efficient with rich data, i.e. when a significant amount of information is available in the data for recovering accurately all the individual parameters. On the contrary, tests and plots can be misleading when the estimates of the individual parameters are greatly shrunk.

We propose to develop new approaches for diagnosing mixed effects models in a general context and derive formal and unbiased statistical tests for testing separately each feature of the model.

3.8. Missing data

The ability to easily collect and gather a large amount of data from different sources can be seen as an opportunity to better understand many processes. It has already led to breakthroughs in several application areas. However, due to the wide heterogeneity of measurements and objectives, these large databases often exhibit an extraordinary high number of missing values. Hence, in addition to scientific questions, such data also present some important methodological and technical challenges for data analyst.

Missing values occur for a variety of reasons: machines that fail, survey participants who do not answer certain questions, destroyed or lost data, dead animals, damaged plants, etc. Missing values are problematic since most statistical methods can not be applied directly on a incomplete data. Many progress have been made to properly handle missing values. However, there are still many challenges that need to be addressed in the future, that are crucial for the users.

- State of arts methods often consider the case of continuous or categorical data whereas real data are very often mixed. The idea is to develop a multiple imputation method based on a specific principal component analysis (PCA) for mixed data. Indeed, PCA has been used with success to predict (impute) the missing values. A very appealing property is the ability of the method to handle very large matrices with large amount of missing entries.
- The asymptotic regime underlying modern data is not any more to consider that the sample size increases but that both number of observations and number of variables are very large. In practice first experiments showed that the coverage properties of confidence areas based on the classical methods to estimate variance with missing values varied widely. The asymptotic method and the bootstrap do well in low-noise setting, but can fail when the noise level gets high or when the number of variables is much greater than the number of rows. On the other hand, the jackknife has good coverage properties for large noisy examples but requires a minimum number of variables to be stable enough.
- Inference with missing values is usually performed under the assumption of "Missing at Random" (MAR) values which means that the probability that a value is missing may depend on the observed data but does not depend on the missing value itself. In real data and in particular in data coming from clinical studies, both "Missing Non at Random" (MNAR) and MAR values occur. Taking into account in a proper way both types of missing values is extremely challenging but is worth investigating since the applications are extremely broad.

It is important to stress that missing data models are part of the general incomplete data models addressed by XPOP. Indeed, models with latent variables (i.e. non observed variables such as random effects in a mixed effects model), models with censored data (e.g. data below some limit of quantification) or models with dropout mechanism (e.g. when a subject in a clinical trial fails to continue in the study) can be seen as missing data models.

4. Application Domains

4.1. Precision medicine and pharmacogenomics

Pharmacogenomics involves using an individual's genome to determine whether or not a particular therapy, or dose of therapy, will be effective. Indeed, people's reaction to a given drug depends on their physiological state and environmental factors, but also to their individual genetic make-up.

Precision medicine is an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person. While some advances in precision medicine have been made, the practice is not currently in use for most diseases.

Currently, in the traditional population approach, inter-individual variability in the reaction to drugs is modeled using covariates such as weight, age, sex, ethnic origin, etc. Genetic polymorphisms susceptible to modify pharmacokinetic or pharmacodynamic parameters are much harder to include, especially as there are millions of possible polymorphisms (and thus covariates) per patient.

The challenge is to determine which genetic covariates are associated to some PKPD parameters and/or implicated in patient responses to a given drug.

Another problem encountered is the dependence of genes, as indeed, gene expression is a highly regulated process. In cases where the explanatory variables (genomic variants) are correlated, Lasso-type methods for model selection are thwarted.

There is therefore a clear need for new methods and algorithms for the estimation, validation and selection of mixed effects models adapted to the problems of genomic medicine.

A target application of this project concerns the lung cancer.

EGFR (Epidermal Growth Factor Receptor) is a cell surface protein that binds to epidermal growth factor. We know that deregulation of the downstream signaling pathway of EGFR is involved in the development of lung cancers and several gene mutations responsible for this deregulation are known.

Our objective is to identify the variants responsible for the disruption of this pathway using a modelling approach. The data that should be available for developing such model are ERK (Extracellular signal–regulated kinases) phosphorylation time series, obtained from different genetic profiles.

The model that we aim to develop will describe the relationship between the parameters of the pathway and the genomic covariates, i.e. the genetic profile. Variants related to the pathway include: variants that modify the affinity binding of ligands to receptors, variants that modify the total amount of protein, variants that affect the catalytic site,...

4.2. Oncology

In cancer, the most dreadful event is the formation of metastases that disseminate tumor cells throughout the organism. Cutaneous melanoma is a cancer, where the primary tumor can easily be removed by surgery. However, this cancer is of poor prognosis; because melanomas metastasize often and rapidly. Many melanomas arise from excessive exposure to mutagenic UV from the sun or sunbeds. As a consequence, the mutational burden of melanomas is generally high

RAC1 encodes a small GTPase that induces cell cycle progression and migration of melanoblasts during embryonic development. Patients with the recurrent P29S mutation of RAC1 have 3-fold increased odds at having regional lymph nodes invaded at the time of diagnosis. RAC1 is unlikely to be a good therapeutic target, since a potential inhibitor that would block its catalytic activity, would also lock it into the active GTP-bound state. This project thus investigates the possibility of targeting the signaling pathway downstream of RAC1.

XPOP is mainly involved in Task 1 of the project: *Identifying deregulations and mutations of the ARP2/3 pathway in melanoma patients*.

Association of over-expression or down-regulation of each marker with poor prognosis in terms of invasion of regional lymph nodes, metastases and survival, will be examined using classical univariate and multivariate analysis. We will then develop specific statistical models for survival analysis in order to associate prognosis factors to each composition of complexes. Indeed, one has to implement the further constraint that each subunit has to be contributed by one of several paralogous subunits. An original method previously developed by XPOP has already been successfully applied to WAVE complex data in breast cancer.

The developed models will be rendered user-friendly though a dedicated Rsoftware package.

This project can represent a significant step forward in precision medicine of the cutaneous melanoma.

4.3. Hemodialysis

Hemodialysis is a process for removing waste and excess water from the blood and is used primarily as an artificial replacement for lost kidney function in people with kidney failure. Side effects caused by removing too much fluid and/or removing fluid too rapidly include low blood pressure, fatigue, chest pains, leg-cramps, nausea and headaches.

Nephrologists must therefore correctly assess the hydration status in chronic hemodialysis patients and consider fluid overload effects when prescribing dialysis, according to a new study.

The fluid overload biomarker, B-type natriuretic peptide (BNP) is an important component of managing patients with kidney disease. Indeed, it is believed that each dialysis patient will have an ideal or "dry" BNP level which will accurately and reproducibly reflect their optimal fluid status.

The objective of this study is to develop a model for the BNP and the hydratation status using individual information (age, sex, ethnicity, systolic blood pressure, BMI, coronary heart disease history, ...).

The impact will be significant if the method succeeds. Indeed, it will be possible for the nephrologists to use this model for monitoring individually each treatment, in order to avoid risks of hypotension (low BNP) or overweight (high BNP).

4.4. Intracellular processes

Significant cell-to-cell heterogeneity is ubiquitously-observed in isogenic cell populations. Cells respond differently to a same stimulation. For example, accounting for such heterogeneity is essential to quantitatively understand why some bacteria survive antibiotic treatments, some cancer cells escape drug-induced suicide, stem cell do not differentiate, or some cells are not infected by pathogens.

The origins of the variability of biological processes and phenotypes are multifarious. Indeed, the observed heterogeneity of cell responses to a common stimulus can originate from differences in cell phenotypes (age, cell size, ribosome and transcription factor concentrations, etc), from spatio-temporal variations of the cell environments and from the intrinsic randomness of biochemical reactions. From systems and synthetic biology perspectives, understanding the exact contributions of these different sources of heterogeneity on the variability of cell responses is a central question.

The main ambition of this project is to propose a paradigm change in the quantitative modelling of cellular processes by shifting from mean-cell models to single-cell and population models. The main contribution of XPOP focuses on methodological developments for mixed-effects model identification in the context of growing cell populations.

- Mixed-effects models usually consider an homogeneous population of independent individuals. This assumption does not hold when the population of cells (i.e. the statistical individuals) consists of several generations of dividing cells. We then need to account for inheritance of single-cell parameters in this population. More precisely, the problem is to attribute the new state and parameter values to newborn cells given (the current estimated values for) the mother.
- The mixed-effects modelling framework corresponds to a strong assumption: differences between cells are static in time (ie, cell-specific parameters have fixed values). However, it is likely that for any given cell, ribosome levels slowly vary across time, since like any other protein, ribosomes are produced in a stochastic manner. We will therefore extend our modelling framework so as to account for the possible random fluctuations of parameter values in individual cells. Extensions based on stochastic differential equations will be investigated.
- Identifiability is a fundamental prerequisite for model identification and is also closely connected to optimal experimental design. We will derive criteria for theoretical identifiability, in which different parameter values lead to non-identical probability distributions, and for structural identifiability, which concerns the algebraic properties of the structural model, i.e. the ODE system. We will then address the problem of practical identifiability, whereby the model may be theoretically identifiable but the design of the experiment may make parameter estimation difficult and imprecise. An interesting problem is whether accounting for lineage effects can help practical identifiability of the parameters of the individuals in presence of measurement and biological noise.

4.5. Population pharmacometrics

Pharmacometrics involves the analysis and interpretation of data produced in pre-clinical and clinical trials. Population pharmacokinetics studies the variability in drug exposure for clinically safe and effective doses by focusing on identification of patient characteristics which significantly affect or are highly correlated with this variability. Disease progress modeling uses mathematical models to describe, explain, investigate and predict the changes in disease status as a function of time. A disease progress model incorporates functions describing natural disease progression and drug action.

The model based drug development (MBDD) approach establishes quantitative targets for each development step and optimizes the design of each study to meet the target. Optimizing study design requires simulations, which in turn require models. In order to arrive at a meaningful design, mechanisms need to be understood and correctly represented in the mathematical model. Furthermore, the model has to be predictive for future studies. This requirement precludes all purely empirical modeling; instead, models have to be mechanistic.

In particular, physiologically based pharmacokinetic models attempt to mathematically transcribe anatomical, physiological, physical, and chemical descriptions of phenomena involved in the ADME (Absorption - Distribution - Metabolism - Elimination) processes. A system of ordinary differential equations for the quantity of substance in each compartment involves parameters representing blood flow, pulmonary ventilation rate, organ volume, etc.

The ability to describe variability in pharmacometrics model is essential. The nonlinear mixed-effects modeling approach does this by combining the structural model component (the ODE system) with a statistical model, describing the distribution of the parameters between subjects and within subjects, as well as quantifying the unexplained or residual variability within subjects.

The objective of XPOP is to develop new methods for models defined by a very large ODE system, a large number of parameters and a large number of covariates. Contributions of XPOP in this domain are mainly methodological and there is no privileged therapeutic application at this stage.

However, it is expected that these new methods will be implemented in software tools, including MONOLIX and Rpackages for practical use.

4.6. Mass spectrometry

One of the main recent developments in analytical chemistry is the rapid democratization of high-resolution mass spectrometers. These instruments produce extremely complex mass spectra, which can include several hundred thousand ions when analyzing complex samples. The analysis of complex matrices (biological, agri-food, cosmetic, pharmaceutical, environmental, etc.) is precisely one of the major analytical challenges of this new century. Academic and industrial researchers are particularly interested in trying to quickly and effectively establish the chemical consequences of an event on a complex matrix. This may include, for example, searching for pesticide degradation products and metabolites in fruits and vegetables, photoproducts of active ingredients in a cosmetic emulsion exposed to UV rays or chlorination products of biocides in hospital effluents. The main difficulty of this type of analysis is based on the high spatial and temporal variability of the samples, which is in addition to the experimental uncertainties inherent in any measurement and requires a large number of samples and analyses to be carried out and computerized data processing (up to 16 million per mass spectrum).

A collaboration between XPOP and the Molecular Chemistry Laboratory (LCM) of the Ecole Polytechnique began in 2018. Our objective is to develop new methods for the statistical analysis of mass spectrometry data.

5. Highlights of the Year

5.1. Highlights of the Year

Version 1.0 of the SPIX software was available in November 2018.

6. New Software and Platforms

6.1. mlxR

KEYWORDS: Simulation - Data visualization - Clinical trial simulator

FUNCTIONAL DESCRIPTION: The models are encoded using the model coding language 'Mlxtran', automatically converted into C++ codes, compiled on the fly and linked to R using the 'Rcpp' package. That allows one to implement very easily complex ODE-based models and complex statistical models, including mixed effects models, for continuous, count, categorical, and time-to-event data.

- Contact: Marc Lavielle
- URL: http://simulx.webpopix.org/

6.2. Rsmlx

R speaks Monolix
KEYWORDS: Data modeling - Nonlinear mixed effects models - Statistical modeling
FUNCTIONAL DESCRIPTION: Among other tasks, 'Rsmlx' provides a powerful tool for automatic PK model building, performs statistical tests for model assessment, bootstrap simulation and likelihood profiling for computing confidence intervals. 'Rsmlx' also proposes several automatic covariate search methods for mixed effects models.

- Partner: Lixoft
- Contact: Marc Lavielle

6.3. SPIX

KEYWORDS: Data modeling - Mass spectrometry - Chemistry
FUNCTIONAL DESCRIPTION: SPIX allows you to - To automatically identify, on the basis of statistical approaches, small but significant differences in spectra measured under different conditions, - Model the kinetics of entities that evolve over time

- Partner: Laboratoire de Chimie Moléculaire - Ecole Polytechnique
- Contact: Marc Lavielle

7. New Results

7.1. Normalizing constants of log-concave densities

We derive explicit bounds for the computation of normalizing constants Z for log-concave densities $\pi = \mathrm{e}^{-U}/Z$ w.r.t. the Lebesgue measure on $\mathbb{R}^d$. Our approach relies on a Gaussian annealing combined with recent and precise bounds on the Unadjusted Langevin Algorithm (Durmus, A. and Moulines, E. (2016). High-dimensional Bayesian inference via the Unadjusted Langevin Algorithm). Polynomial bounds in the dimension d are obtained with an exponent that depends on the assumptions made on U. The algorithm also provides a theoretically grounded choice of the annealing sequence of variances. A numerical experiment supports our findings. Results of independent interest on the mean squared error of the empirical average of locally Lipschitz functions are established.

7.2. The Tamed Unadjusted Langevin Algorithm

We consider the problem of sampling from a probability measure π having a density on $\mathbb{R}^d$ known up to a normalizing constant, $x \to \mathrm{e}^{-U(x)}/Z$. The Euler discretization of the Langevin stochastic differential equation (SDE) is known to be unstable in a precise sense, when the potential U is superlinear. Based on previous works on the taming of superlinear drift coefficients for SDEs, we introduce the Tamed Unadjusted Langevin Algorithm (TULA) and obtain non-asymptotic bounds in V-total variation norm and Wasserstein distance of order 2 between the iterates of TULA and π, as well as weak error bounds. Numerical experiments support our findings.

7.3. Development and performance of npde for the evaluation of time-to-event models

Normalised prediction distribution errors (npde) are used to graphically and statistically evaluate mixed-effect models for continuous responses. Our aim was to extend npde to time-to-event (TTE) models and evaluate their performance. We extended npde to TTE models using imputations to take into account censoring. We then evaluated their performance in terms of type I error and power to detect model misspecifications for TTE data by means of a simulation study with different sample sizes. Type I error was found to be close to the expected 5% significance level for all sample sizes tested. The npde were able to detect misspecifications in the baseline hazard as well as in the link between the longitudinal variable and the survival function. The ability to detect model misspecifications increased as the difference in the shape of the survival function became more apparent. As expected, the power also increased as the sample size increased. Imputing the censored events tended to decrease the percentage of rejections.

7.4. Low-rank Interaction with Sparse Additive Effects Model for Large Data Frames

Many applications of machine learning involve the analysis of large data frames-matrices collecting heterogeneous measurements (binary, numerical, counts, etc.) across samples-with missing values. Low-rank models are popular in this framework for tasks such as visualization, clustering and missing value imputation. Yet, available methods with statistical guarantees and efficient optimization do not allow explicit modeling of main additive effects such as row and column, or covariate effects. We introduced a low-rank interaction and sparse additive effects (LORIS) model which combines matrix regression on a dictionary and low-rank design, to estimate main effects and interactions simultaneously. We provide statistical guarantees in the form of upper bounds on the estimation error of both components. Then, we introduced a mixed coordinate gradient descent (MCGD) method which provably converges sub-linearly to an optimal solution and is computationally efficient for large scale data sets. Simulated and survey data showed that the method has a clear advantage over current practices, which consist in dealing separately with additive effects in a preprocessing step.

7.5. Diffusion approximations and control variates for MCMC

A new methodology was developed for the construction of control variates to reduce the variance of additive functionals of Markov Chain Monte Carlo (MCMC) samplers. Our control variates are defined as linear combinations of functions whose coefficients are obtained by minimizing a proxy for the asymptotic variance. The construction is theoretically justified by two new results. We first show that the asymptotic variances of some well-known MCMC algorithms, including the Random Walk Metropolis and the (Metropolis) Unadjusted/Adjusted Langevin Algorithm, are close to the asymptotic variance of the Langevin diffusion. Second, we provide an explicit representation of the optimal coefficients minimizing the asymptotic variance of the Langevin diffusion. Several examples of Bayesian inference problems demonstrate that the corresponding reduction in the variance is significant, and that in some cases it can be dramatic.

7.6. Density estimation for random walks in random environment

We consider the problem of non-parametric density estimation of a random environment from the observation of a single trajectory of a random walk in this environment. We first construct a density estimator using the beta-moments. We then show that the Goldenshluger-Lepski method can be used to select the beta-moment. We prove nonasymptotic bounds for the supremum norm of these estimators for both the recurrent and the transient to the right case. A simulation study supports our theoretical findings.

7.7. Imputation of mixed data with multilevel singular value decomposition

Statistical analysis of large data sets offers new opportunities to better understand many processes. Yet, data accumulation often implies relaxing acquisition procedures or compounding diverse sources. As a consequence, such data sets often contain mixed data, i.e. both quantitative and qualitative and many missing values. Furthermore, aggregated data present a natural multilevel structure, where individuals or samples are nested within different sites, such as countries or hospitals. Imputation of multilevel data has therefore drawn some attention recently, but current solutions are not designed to handle mixed data, and suffer from important drawbacks such as their computational cost. In this article, we propose a single imputation method for multilevel data, which can be used to complete either quantitative, categorical or mixed data. The method is based on multilevel singular value decomposition (SVD), which consists in decomposing the variability of the data into two components, the between and within groups variability, and performing SVD on both parts. We show on a simulation study that in comparison to competitors, the method has the great advantages of handling data sets of various size, and being computationally faster. Furthermore, it is the first so far to handle mixed data. We apply the method to impute a medical data set resulting from the aggregation of several data sets coming from different hospitals. This application falls in the framework of a larger project on Trauma patients. To overcome obstacles associated to the aggregation of medical data, we turn to distributed computation. The method is implemented in an R package

7.8. Logistic Regression with Missing Covariates – Parameter Estimation, Model Selection and Prediction

Logistic regression is a common classification method in supervised learning. Surprisingly , there are very few solutions for performing it and selecting variables in the presence of missing values. We develop a complete approach, including the estimation of parameters and variance of estimators, derivation of confidence intervals and a model selection procedure, for cases where the missing values can be anywhere in covariates. By well organizing different patterns of missingness in each observation , we propose a stochastic approximation version of the EM algorithm based on Metropolis-Hasting sampling, to perform statistical inference for logistic regression with incomplete data. We also tackle the problem of prediction for a new individual with missing values, which is never addressed. The methodology is computationally efficient, and its good coverage and variable selection properties are demonstrated in a simulation study where we contrast its performances to other methods. For instance, the popular multiple imputation by chained equation can lead to biased estimates while our method is unbiased. The method was applied on a dataset of severely traumatized patients from Paris hospitals to predict the occurrence of hemorrhagic shock, a leading cause of early preventable death in severe trauma cases. The aim is to consolidate the current red flag procedure, a binary alert identifying patients with a high risk of severe hemorrhage. The methodology is implemented in the R package misaem.

7.9. A fast Stochastic Approximation of the EM algorithm for nonlinear mixed effects models

The ability to generate samples of the random effects from their conditional distributions is fundamental for inference in mixed effects models. Random walk Metropolis is widely used to perform such sampling, but this method is known to converge slowly for high dimensional problems, or when the joint structure of the distributions to sample is spatially heterogeneous. We propose an independent Metropolis-Hastings (MH) algorithm based on a multidimensional Gaussian proposal that takes into account the joint conditional distribution of the random effects and does not require any tuning. Indeed, this distribution is automatically obtained thanks to a Laplace approximation of the incomplete data model. We show that such approximation is equivalent to linearizing the structural model in the case of continuous data. Numerical experiments based on simulated and real data demonstrate the good performance of the proposed methods. In particular, we show that the suggested MH algorithm can be efficiently combined with a stochastic approximation version of the EM algorithm for maximum likelihood estimation in nonlinear mixed effects models.

7.10. Incomplete graphical model inference via latent tree aggregation

Graphical network inference is used in many fields such as genomics or ecology to infer the conditional independence structure between variables, from measurements of gene expression or species abundances for instance. In many practical cases, not all variables involved in the network have been observed, and the samples are actually drawn from a distribution where some variables have been marginalized out. This challenges the sparsity assumption commonly made in graphical model inference, since marginalization yields locally dense structures, even when the original network is sparse. We developed a procedure for inferring Gaussian graphical models when some variables are unobserved, that accounts both for the influence of missing variables and the low density of the original network. Our model is based on the aggregation of spanning trees, and the estimation procedure on the Expectation-Maximization algorithm. We treat the graph structure and the unobserved nodes as missing variables and compute posterior probabilities of edge appearance. To provide a complete methodology, we also propose several model selection criteria to estimate the number of missing nodes. A simulation study and an illustration flow cytometry data reveal that our method has favorable edge detection properties compared to existing graph inference techniques. The methods are implemented in an R package.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Contract with Dassault Systèmes

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

Mixed-Effects Models of Intracellular Processes: Methods, Tools and Applications (MEMIP)

Coordinator: Gregory Batt (InBio Inria team)

Other partners: InBio and IBIS Inria teams, Laboratoire Matière et Systèmes Complexes (UMR 7057; CNRS and Paris Diderot Univ.)

9.1.2. Institut National du Cancer (INCa)

Targeting Rac-dependent actin polymerization in cutaneous melanoma - Institut National du Cancer

Coordinator: Alexis Gautreau (Ecole Polytechnique)

Other partners: Laboratoire de Biochimie (Polytechnique), Institut Curie, INSERM.

9.2. International Initiatives

9.2.1. International Initiatives

SaSMoTiDep

Title: Statistical and Stochastic modeling for time-dependent data

International Partners (Institution - Laboratory - Researcher):

Universidad de Valparaiso (Chile) - Centro de Investigación y Modelamiento de Fenómenos Aleatorios Valparaíso (CIMFAV) - Cristian Meza Becerra

Universidad Nacional de Colombia (Colombia) - Department of Statistics - Viswanathan Arunachalam

Duration: 01/01/2018 - 31/12/2019

Start year: 2018

See also: https://sasmotidep.uv.cl

In many applications, multiple measurements are made on one or several experimental units over a period of time. Such data could be called time-dependent data. From a statistical point of view, if we consider only one experimental unit, we can use a time series analysis. In the other hand, if we consider experimental designs (or observational studies) for several experimental units (or subjects) where each subject is measured at several points in time, we can use the term longitudinal data. In this project, we propose to study several statistical and stochastic models for repeated measures using parametric and non-parametric approaches. In particular, we will study the inference in complex mixed effects models, we will propose novel segmentation models for multiple series, non-parametric methods in dependent models and stochastic models. We will apply these methods to real data from several fields as biometrics, reliability, population dynamics and finance.

9.3. International Research Visitors

9.3.1. Visits of International Scientists

Ricardo Rios, Universidad Central de Venezuela, Caracas: September 2018.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- Winter School New Frontiers in high-dimensional probability and statistics, Moscow, February 23-24, 2018
- Workshop "Structural Inference in Inference in High-Dimensional Models", Moscow, 5-8 September 2018

10.1.2. Scientific Events Selection

10.1.2.1. Member of the Conference Program Committees

- NIPS 2018
- COLT 2018
- useR!2018

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Stochastic Processes and their Applications
- Journal of Statistical Planning and Inference
- Journal of Computational and Graphical Statistics

10.1.4. Invited Talks

- Analysis of adaptive stochastic gradient and MCMC algorithms, June 2018
- Langevin MCMC: theory and methods, Singapour, June 2018

- MCQMC 2018: Langevin Monte Carlo: theory and methods, Rennes, July 2018
- Langevin MCMC: theory and methods, Singapour, August 2018
- Huawei European Innovation Day, "Toward Bayesian Machine Learning", Roma, November 2018
- Iberoamerican Pharmacometrics Network Congress 2018, Guadalajara, Mexico.
- "Tests d'hypothèse et algorithmes de construction de modèle pour les modèles à effets mixtes", Toulouse, September 2018.
- New algorithms and tools for population model building and evaluation, Valparaiso, Chile, December 2018.

10.1.5. Leadership within the Scientific Community

- Eric Moulines is in charge of the academic supervision of the International Laboratory of Stochastic Algorithms and High-dimensional inference, National Research University, Higher School of Economics, funded by the Russian Academic Excellence Project.
- Eric Moulines is is associate researcher of the Alan Turing Institute
- Eric Moulines is elected member of the French Académie des Sciences.

10.1.6. Scientific Expertise

- Marc Lavielle is member of the Scientific Committee of the High Council for Biotechnologies.
- Marc Lavielle is member of the evaluation committee of the International Center for Mathematics (CIMAT), Guanajuato, Mexico.
- Eric Moulines is member of the award committee of foundation "Charles Defforey".

10.1.7. Research administration

- Marc Lavielle is member of the Scientific Programming Committee (CPS) of the Institute Henri Poincaré (IHP).
- Eric Moulines is a board member of the Institut de Convergence DataIA.
- Julie Josse is elected member of the R foundation.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master : Julie Josse, Statistics with R, 48, M2, X-HEC

Master : Eric Moulines, Regression models, 36, M2, X-HEC

Engineering School : Eric Moulines, Statistics, 36, 2A, X

Engineering School : Eric Moulines, Markov Chains, 36, 3A, X

Engineering School : Erwan Le Pennec, Statistics, 36, 2A, X

Engineering School : Erwan Le Pennec, Statistical Learning, 36, 3A, X

Engineering School : Marc Lavielle, Statistics in Action, 48, 3A, X

10.2.2. Supervision

PhD in progress : Nicolas Brosse, September 2016, Eric Moulines

PhD in progress : Geneviève Robin, September 2016, Julie Josse and Eric Moulines

PhD in progress : Belhal Karimi, October 2016, Marc Lavielle and Eric Moulines

PhD in progress : Marine Zulian, October 2016, Marc Lavielle

PhD in progress : Wei Jiang , October 2017, Julie Josse and Marc Lavielle

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

Eric Moulines is member of the Evaluation Committee of Inria.

10.3.2. Creation of media or tools for science outreach

Marc Lavielle developed and maintains the learning platform Statistics in Action. The purpose of this online learning platform is to show how statistics (and biostatistics) may be efficiently used in practice using R. It is specifically geared towards teaching statistical modelling concepts and applications for self-study. Indeed, most of the available teaching material tends to be quite "static" while statistical modelling is very much subject to "learning by doing".

Julie Josse (with Nicholas Tierney and Nathalie Vialaneix) developed and maintains the website R-miss-tastic, A resource website on missing data.

11. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] F. ALLA, N. AGRINIER, M. LAVIELLE, P. ROSSIGNOL, D. GONTHIER, J.-M. BOIVIN, F. ZANNAD. *Impact of the interruption of a large heart failure regional disease management program on hospital admission rate: a population-based study*, in "European Journal of Heart Failure", 2018, https://hal.inria.fr/hal-01727632

[2] N. BROSSE, A. DURMUS, É. MOULINES. *Normalizing constants of log-concave densities*, in "Electronic journal of statistics ", March 2018, vol. 12, n^{o} 1, p. 851-889 [*DOI :* 10.1214/18-EJS1411], https://hal.inria.fr/hal-01648666

[3] N. BROSSE, A. DURMUS, É. MOULINES, S. SABANIS. *The Tamed Unadjusted Langevin Algorithm*, in "Stochastic Processes and their Applications", October 2018 [*DOI :* 10.1016/J.SPA.2018.10.002], https://hal.inria.fr/hal-01648667

[4] M. CEROU, M. LAVIELLE, K. BRENDEL, M. CHENEL, E. COMETS. *Development and performance of npde for the evaluation of time-to-event models*, in "Pharmaceutical Research", January 2018, vol. 35, n^{o} 2, 30 [*DOI :* 10.1007/s11095-017-2291-3], https://www.hal.inserm.fr/inserm-01695500

[5] M. LAVIELLE. *Pharmacometrics Models with Hidden Markovian Dynamics*, in "Journal of Pharmacokinetics and Pharmacodynamics", 2018, vol. 45, n^{o} 1, p. 91–105 [*DOI :* 10.1007/S10928-017-9541-1], https://hal.inria.fr/hal-01665722

[6] L. LÊ, M. BERGE, A. TFAYLI, A. BAILLET-GUFFROY, P. PROGNON, A. DOWEK, E. CAUDRON. *Quantification of gemcitabine intravenous drugs by direct measurement in chemotherapy plastic bags using a handheld Raman spectrometer*, in "Talanta", May 2019, vol. 196, p. 376-380, https://hal.archives-ouvertes.fr/hal-01970020

[7] L. M. M. LÊ, M. BERGE, A. TFAYLI, J. ZHOU, P. PROGNON, A. BAILLET-GUFFROY, E. CAUDRON. *Rapid discrimination and quantification analysis of five antineoplastic drugs in aqueous solutions using Raman spectroscopy*, in "European Journal of Pharmaceutical Sciences", January 2018, vol. 111, p. 158-166, https://hal.archives-ouvertes.fr/hal-01970012

[8] L. M. M. LÊ, B. KÉGL, A. GRAMFORT, C. MARINI, D. NGUYEN, M. CHERTI, S. TFAILI, A. TFAYLI, A. BAILLET-GUFFROY, P. PROGNON, P. CHAMINADE, E. CAUDRON. *Optimization of classification and regression analysis of four monoclonal antibodies from Raman spectra using collaborative machine learning approach*, in "Talanta", July 2018, vol. 184, p. 260-265, https://hal.archives-ouvertes.fr/hal-01969999

[9] L. LÊ, A. TFAYLI, P. PROGNON, E. CAUDRON.*Discriminative and Quantitative Analysis of Antineoplastic Taxane Drugs Using a Handheld Raman Spectrometer*, in "BioMed Research International ", July 2018, vol. 2018, p. 1-7, https://hal.archives-ouvertes.fr/hal-01969980

International Conferences with Proceedings

[10] N. BROSSE, A. DURMUS, É. MOULINES.*The promises and pitfalls of Stochastic Gradient Langevin Dynamics*, in "NeurIPS 2018 (Advances in Neural Information Processing Systems 2018)", Montreal, France, December 2018, https://hal.inria.fr/hal-01934291

[11] G. ROBIN, H.-T. WAI, J. JOSSE, O. KLOPP, É. MOULINES.*Low-rank Interaction with Sparse Additive Effects Model for Large Data Frames*, in "32nd Conference on Neural Information Processing Systems (NeurIPS 2018)", Montréal, Canada, December 2018, https://arxiv.org/abs/1812.08398 , https://hal.archives-ouvertes.fr/hal-01959188

Conferences without Proceedings

[12] C. BIERNACKI, G. CELEUX, J. JOSSE, F. LAPORTE.*Dealing with missing data in model-based clustering through a MNAR model*, in "CMStatistics 2018 - 11th International Conference of the ERCIM WG on Computational and Methodological Statistics", Pise, Italy, December 2018, https://hal.archives-ouvertes.fr/hal-01949120

Other Publications

[13] N. BROSSE, A. DURMUS, S. MEYN, É. MOULINES.*Diffusion approximations and control variates for MCMC*, November 2018, working paper or preprint, https://hal.inria.fr/hal-01934316

[14] A. HAVET, M. LERASLE, É. MOULINES.*Density estimation for RWRE*, June 2018, https://arxiv.org/abs/1806.05839 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01815990

[15] F. HUSSON, J. JOSSE, B. NARASIMHAN, G. ROBIN.*Imputation of mixed data with multilevel singular value decomposition*, April 2018, https://arxiv.org/abs/1804.11087 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01781291

[16] W. JIANG, J. JOSSE, M. LAVIELLE.*Logistic Regression with Missing Covariates – Parameter Estimation, Model Selection and Prediction*, December 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01958835

[17] B. KARIMI, M. LAVIELLE.*Efficient Metropolis-Hastings sampling for nonlinear mixed effects models*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01958247

[18] B. KARIMI, M. LAVIELLE, É. MOULINES.*f-SAEM: A fast Stochastic Approximation of the EM algorithm for nonlinear mixed effects models*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01958248

[19] L. MONTUELLE, E. LE PENNEC.*PAC-Bayesian aggregation of affine estimators*, January 2018, https://arxiv.org/abs/1410.0661 - working paper or preprint, https://hal.inria.fr/hal-01070805

[20] G. ROBIN, C. AMBROISE, S. S. ROBIN.*Incomplete graphical model inference via latent tree aggregation*, January 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01686841

[21] A. SPORTISSE, C. BOYER, J. JOSSE.*Imputation and low-rank estimation with Missing Non At Random data*, January 2019, https://arxiv.org/abs/1812.11409 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01964720